

**Astrophysics
and the
Many-Body Problem**

BRANDEIS UNIVERSITY SUMMER INSTITUTE LECTURES IN THEORETICAL PHYSICS

K. W. Ford, *Editor*

1960 Lectures

C. Møller • P. T. Matthews • J. Schwinger • N. Fukuda •
J. J. Sakurai

1961 Lectures

Vol. 1

R. J. Eden • J. C. Polkinghorne • G. Källén • J. J. Sakurai

Vol. 2

M. E. Rose • E. C. G. Sudarshan

1962 Lectures

Vol. 1—Elementary Particle Physics and Field Theory

T. Fulton • G. Källén • J. D. Jackson • C. Fronsdal

Vol. 2—Astrophysics and the Many-Body Problem

E. N. Parker • J. S. Goldstein • A. A. Maradudin •
V. Ambegaokar

Vol. 3—Statistical Physics

G. E. Uhlenbeck • N. Rosenzweig • A. J. F. Siegert •
E. T. Jaynes • S. Fujita

Brandeis Summer Institute 1962

**ASTROPHYSICS
AND
THE MANY-BODY PROBLEM**

2

**E. N. Parker
J. S. Goldstein
A. A. Maradudin
V. Ambegaokar**

Notes by
**M. Turoff
A. Skalafuris
B. Bergerson**

W. A. BENJAMIN, INC.

1963

New York

Amsterdam

IIA Lib.,

ASTROPHYSICS AND THE MANY-BODY PROBLEM
Volume 2 of the 1962 Brandeis Summer Institute
Lectures in Theoretical Physics

Copyright © 1963 by W. A. Benjamin, Inc.
All rights reserved

Library of Congress Catalog Card Number 61-18591
Manufactured in the United States of America

*Final manuscript was received on December 7, 1962; this volume was
published August 30, 1963*

W. A. BENJAMIN, INC.
2465 Broadway, New York 25, New York

Foreword

It is now an established tradition of the Brandeis Summer Institute in Theoretical Physics to have lecturers who present a systematic account of recent research in various fields of theoretical physics. The lecture notes have also become part of this tradition, and, although these are sometimes but a first approximation to the spoken lecture, they may serve to bring these much-needed expositions to the wider audience of physicists who may aspire to contribute to these fields.

I should like to take this opportunity to thank all those whose participation in the Institute during the summer of 1962 helped maintain these traditions. Particular words of appreciation are due the National Science Foundation, for its indispensable financial support, and Professor Kenneth Ford, who graciously carried the responsibility for getting the notes ready for publication.

In this volume, the notes of Professor Goldstein and Dr. Maradudin have been prepared by the lecturers; Professor Parker and Professor Ambegaokar have kindly checked over the notes based on their lectures.

DAVID L. FALKOFF

Co-Director of the 1962 Institute

Contents

Foreword	v
Plasma Dynamics in Astrophysics and Geophysics	1
<i>E. N. Parker, University of Chicago</i> (Notes by M. Turoff and A. Skalafuris)	
Stellar Evolution	83
<i>J. S. Goldstein, Brandeis University</i> (Notes by the lecturer)	
Dynamic Theory of Imperfect Crystals	107
<i>A. A. Maradudin, Westinghouse Laboratories, Pittsburgh</i> (Notes by the lecturer)	
Green's Functions in Many-Body Problems	321
<i>V. Ambegaokar, Cornell University</i> (Notes by B. Bergerson)	

**Astrophysics
and the
Many-Body Problem**

PLASMA PHYSICS AND ASTROPHYSICS

E. N. Parker

Enrico Fermi Institute for Nuclear Studies

University of Chicago

CONTENTS

1. Notation	4
2. Introduction	7
3. Equations of Motion for a Conducting Fluid	7
4. Collisionless Plasma	15
5. The Boltzmann Equation for Collisionless Plasma	20
6. Dynamics of an Ionized Gas	26
7. The Granulation of the Sun	34
8. Acoustical Transport of Coronal Energy	41
9. Acoustical and Internal-Gravity Waves in the Solar Atmosphere	45
10. The Heating of the Solar Corona by Means of Hydromagnetic Waves	49
11. The Statics of a Stellar Atmosphere at One Million Degrees	53
12. Hydrodynamic Expansion of the Solar Corona	57
13. Stellar Winds and Solar Phenomena	64
Interplanetary Conditions	67
Expansion of the Active Corona	69
Bibliography and Notes, Secs. 1 to 6	77
References	80

1. NOTATION

a	Speed of sound
A	An area
B	Magnetic-field intensity
b	Average radius of pulsating sphere
C	Specific heat
C	Alfvén velocity
c	Light velocity
D	Radius of gyration
\mathcal{E}	Energy radiated per second per cubic centimeter
E	Electric-field intensity
F	External force on a particle
$f(x_i, u_j, t)$	Differential phase-space particle-distribution function
f	Number of degrees of freedom
G	Gravitational constant
g	Gravitational-field intensity
(i, j, k, l)	Vector component subscripts used alternately with (x, y, z)
H	Scale height
J	Current density
\mathbf{k}	Propagation vector (an inverse length)
k	Boltzmann constant
K	Thermometric conductivity
\mathcal{K}	Curvature of the lines of force
L, l	A characteristic scale or length of the magnetic field
M_{ij}	Maxwell stress tensor
M	Mass of a star
M_\odot	Mass of the sun
m	Particle mass (electron or proton)

\hat{n}	Normal unit vector
n	An integer index
N	Particle density
P	Power radiated
$P(\eta)$	Similarity variable for the pressure
p, p_{ij}	Scalar or tensor pressure
q	Algebraic electric charge (\pm)
Q	Heat transferred to the system
Q_{ijk}	Heat-transfer tensor
R	Reynolds' number
R^m	Magnetic Reynolds' number
\mathcal{R}	Rayleigh's number
R_\odot	Radius of reference in a star
S_{ij}	General fluid stress tensor
T	Temperature
t	Time
u	Microscopic particle velocity
u_d	Particle drift velocity
$U(\eta)$	Similarity variable simulating fluid velocity behind the shock wave
v	Macroscopic fluid velocity
V	Shock-wave velocity
V	Velocity of solar wind in units of thermal velocity
x	Measure in the x direction
x_i	Component of the position vector
α	Coefficient of thermal expansion
α	Polytropic index
β	Temperature gradient
γ	Ratio of specific heats

δ_{ij}	Kronecker delta
ϵ_{ijk}	Asymmetric skew tensor (+) for odd, (-) for even permutation of indices
ζ	Eigenvalue parameter
η	Similarity variable
θ	Angle of trajectory relative to magnetic field
ϕ	Azimuthal angle in spherical coordinate system
κ	Coefficient of thermal conductivity
λ	Wavelength
μ	Coefficient of viscosity
ρ	Mass density
σ	Electrical conductivity
τ	Characteristic time scale
φ	Number of particles per cubic centimeter per unit solid angle
Φ	Hydrodynamic velocity potential
ψ	Number density of particles per unit pitch angle
Ψ	Gravitational potential
Ω	Cyclotron frequency
ω	Angular velocity
\perp	(subscript) Perpendicular to magnetic field
\parallel	(subscript) Parallel to magnetic field

2. INTRODUCTION

Laboratory plasma physics differs from the plasma physics of astrophysics in that it deals with many more degrees of freedom than usually need to be considered in astrophysical problems. In stellar structure the systems are more gross. The internal effects that dominate laboratory generated plasmas are generally smeared out in astrophysical applications. Thus only gross features of plasmas need be used in approaching most astrophysical dynamical phenomena to the present degree of observation.

Our purpose will be, therefore, to develop the general equations of motion of a gross plasma (ignoring the finer structure) and show how these equations can be applied to some astrophysical phenomena. We shall develop our equations for two basic ideal situations: (1) in which we shall assume no collisions take place among the particles of the gas; and in the other extreme (2) in which the collision times are very small and the conductivity high, i. e., a perfectly conducting classical fluid. Several standard works on the subject are listed in the bibliography at the end of these lectures.

It will be shown that these two extreme idealizations lead to similar results in systems of large scale (astrophysical size). The implication that the intermediate region bounded by our two assumptions can be characterized by either assumption is not valid. However, as a practical consideration, stellar structure in many cases is such that as one progresses out from a star one passes very quickly from one region to the next with practically no intermediate region.

3. EQUATIONS OF MOTION FOR A CONDUCTING FLUID

We shall use the following nomenclature:

$\rho(x, t)$	density
$p(x, t)$	pressure
$v_i(x, t)$	velocity
$B_i(x, t)$	magnetic field
$E_i(x, t)$	electric field
$g_i = - \frac{\partial \psi}{\partial x_i}$	gravitational field (the gradient of some potential)
M_{ij}	Maxwell stress tensor
ν_{ij}	viscosity tensor

The first equation we may write down is the conservation of mass:

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial X_i} (\rho v_i) = 0 \quad (3.1)$$

The total derivative defined by

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \frac{dX_i}{dt} \frac{\partial}{\partial X_i} = \frac{\partial}{\partial t} + v_i \frac{\partial}{\partial X_i}$$

allows us to rewrite (3.1) in the following form:

$$\frac{1}{\rho} \frac{d\rho}{dt} + \frac{\partial v_i}{\partial X_i} = 0 \quad (3.2)$$

The momentum equation is

$$\rho \frac{dv_i}{dt} = - \frac{\partial p}{\partial X_i} + F_i \quad (3.2)$$

where

$$F_i = \rho g_i + \frac{\partial M_{ij}}{\partial X_j} + \frac{\partial v_{ij}}{\partial X_j}$$

The Maxwell stress tensor contains two parts, the magnetic- and electric-field contributions:

$$M_{ij}^m = - \delta_{ij} \frac{B^2}{8\pi} + \frac{B_i B_j}{4\pi}$$

$$M_{ij}^e = - \delta_{ij} \frac{E^2}{8\pi} + \frac{E_i E_j}{4\pi}$$

In a highly conducting fluid the electric part is essentially zero, since infinite conductivity does not permit any large charge separation to occur.

The general form of the viscous stress tensor is the following:

$$\nu_{ij} = A \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) + B \delta_{ij} \frac{\partial v_k}{\partial x_k}$$

However, if we require that the viscous stresses vanish upon the contraction of the indices, then the following condition results:

$$\nu_{ij} = \mu \left[\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) - \frac{2}{3} \delta_{ij} \frac{\partial v_k}{\partial x_k} \right] \quad (3.5)$$

where μ is the coefficient of viscosity.

By using Eqs. (3.2) and (3.3), we may form the following equation:

$$\frac{\partial(\rho v_i)}{\partial t} = - \frac{\partial p}{\partial x_j} + F_i - \frac{\partial(\rho v_i v_j)}{\partial x_j}$$

This may be rewritten

$$\frac{\partial(\rho v_i)}{\partial t} = F_i - \frac{\partial S_{ij}}{\partial x_j} \quad (3.6)$$

where

$$S_{ij} = \delta_{ij} P + \rho v_i v_j \quad (3.7)$$

As we see (3.7) has the form of a stress tensor. The $\rho v_i v_j$ part corresponds to a compressive stress along the flow lines which could lead to turbulence at high velocity. If we make a dimensional comparison

of this term with the viscous force, we arrive at the quantity known as the Reynolds number. Since the viscous force opposes the buckling of the compression stress, the Reynolds number gives us a measure for the tendency toward buckling of the flow lines and turbulent motion:

$$R \sim \frac{\rho v^2 / l}{\mu v / l^2} = \frac{l v \rho}{\mu} \quad (3.8)$$

where l is the scale of the system. When l is sufficiently large the tendency to buckle overcomes the viscous damping and turbulent motion results.

The energy equation for the fluid is

$$\rho C \frac{dT}{dt} = Q + \frac{\partial}{\partial X_i} \left(\kappa \frac{\partial T}{\partial X_i} \right) - p \frac{\partial v_i}{\partial X_i} \quad (3.9)$$

where κ is the coefficient of thermal conductivity, Q represents any heat source, ρC is the specific heat which is equal to $(f/2) Nk$, and f is the number of degrees of freedom; k is the Boltzmann constant, 1.38×10^{-16} erg/ $^{\circ}K$, and N is the number of particles per unit volume. In many problems it may be assumed that (1) Q is zero and (2) the thermal conductivity (κ) is zero. Under these assumptions we may write from (3.9)

$$\frac{f}{2} \frac{\partial T}{\partial t} = -T \frac{\partial v_i}{\partial X_i}$$

where f is the number of degrees of freedom of the molecules. We also have from the gas law that

$$\frac{\partial p}{\partial t} = Nk \frac{\partial T}{\partial t} + kT \frac{\partial N}{\partial t}$$

By using assumption (2) and $\rho = Nm$, we have

$$\frac{\partial p}{\partial t} = \left(1 + \frac{2}{f}\right) \frac{kT}{m} \frac{\partial \rho}{\partial t}$$

or

$$\frac{\partial p}{\partial t} = a^2 \frac{\partial \rho}{\partial t} \quad a^2 = \gamma p / \rho, \quad \gamma = 1 + (2/f) \quad (3.10)$$

and a is the speed of sound.

In the frame of reference moving locally with the conducting fluid velocity, the simple scalar Ohm's law is

$$\bar{J} = \sigma \bar{E}'$$

The field E' in the moving frame is related to the fields \bar{E} and \bar{B} in the fixed frame by $\bar{E}' = \bar{E} + (\bar{v}/c) \times \bar{B}$ neglecting terms $\mathcal{O}(v^2/c^2)$. Thus

$$J = \sigma \left(\bar{E} + \frac{\bar{v} \times \bar{B}}{c} \right)$$

or

$$\bar{E} = \frac{\bar{J}}{\sigma} - \frac{\bar{v}}{c} \times \bar{B} \quad (3.11)$$

Maxwell's equations provide us with an equation relating \bar{J} and \bar{B} ,

$$4\pi\bar{J} + \frac{\partial \bar{E}}{\partial t} = c \nabla \times \bar{B} \quad (3.12)$$

In the highly conducting fluid under consideration, the time derivative of the electric field may be ignored since it is $\mathcal{O}(v^2/c^2)$ which is small compared to the other terms (to be shown presently). Hence we have from (3.11) and (3.12)

$$\bar{E} = -\frac{c}{4\pi\sigma} \nabla \times \bar{B} - \frac{\bar{v}}{c} \times \bar{B} \quad (3.13)$$

Then from the other Maxwell equation

$$\frac{\partial \bar{B}}{\partial t} = -c \nabla \times \bar{B} \quad (3.14)$$

we have

$$\frac{\partial \bar{B}}{\partial t} = \nabla \times (\nabla \times \bar{B}) + \frac{c^2}{4\pi\sigma} \nabla^2 \bar{B} \quad (3.15)$$

which is the equation giving the electromagnetic effect of \bar{v} on \bar{B} . If σ is small we obtain a diffusion equation. On the other hand, if we examine the dimensionless ratio formed by the last two terms and require that the diffusion term $\nabla^2 \bar{B}$ be negligible, we are led to the following inequality:

$$\frac{v}{l} \gg \frac{c^2}{4\pi\sigma l^2}$$

The quantity

$$R^m = \frac{4\pi\sigma v l}{c} \quad (3.16)$$

is the magnetic Reynolds number and is a measure of the conductivity of a fluid. The larger R^m is, the more negligible is the diffusion term in (3.15). The equation for the conductivity of fully ionized hydrogen is the following:

$$\sigma \approx 10^7 T^{3/2} \text{ (per sec)}$$

Taking a distance of 1000 km for the thickness of the chromosphere of the sun, a gas velocity of 1 km/sec and a temperature of $10^4 K^{\circ}$ we have roughly:

$$R^m \approx 4\pi \frac{10^{13} 10^8 10^5}{10^{21}} = 10^6$$

So even though σ for fully ionized hydrogen is only 10^{13} at this temperature compared to copper at 10^{17} in the laboratory, we may still consider this to be a highly conducting fluid to very good approximation.

For an infinitely conducting fluid the lines of force are transported bodily by the fluid, the lines remaining permanently connected. This can be shown in the following way: if we consider the change of the magnetic flux in a time δt through an arbitrary fixed area enclosed by a contour \mathcal{L} , then from (3.15) we can write

$$\delta t \frac{\partial}{\partial t} \int \mathbf{B} \cdot d\mathbf{S} = \int d\mathbf{S} \cdot \nabla \times (\nabla \times \mathbf{B}) \delta t$$

From Stokes' theorem the following line integral is obtained:

$$\begin{aligned} \int d\mathbf{S} \cdot \nabla \times (\nabla \times \mathbf{B}) \delta t &= \delta t \oint d\vec{\mathcal{L}} \cdot \nabla \times \mathbf{B} \\ &= \oint \mathbf{B} \cdot \nabla \times \delta \mathcal{L} \delta t \end{aligned}$$

On the other hand, if the enclosing contour were moving with the fluid, then the displacement $\delta \mathbf{s}$ of an element of the contour in a time δt is

$$\delta \mathbf{s} = \mathbf{v} \delta t$$

Thus we may write that the change of flux through the area enclosed by the fixed contour is

$$\delta t \frac{\partial}{\partial t} \int \mathbf{B} \cdot d\mathbf{S} = \int \mathbf{B} \cdot \delta \mathbf{s} \times d\vec{\mathcal{L}}$$

Now $\delta \mathbf{s} \times d\vec{\mathcal{L}}$ is an element of area enclosed between the fixed contour and the contour moving with the fluid after they have been separating for a time δt . The integral of $\mathbf{B} \cdot \delta \mathbf{s} \times d\vec{\mathcal{L}}$ represents the flux through this area enclosed by the contours. The equation tells us that if the flux through the fixed contour has changed, an equivalent amount of flux now passes through the area between the fixed and the moving contour, i. e., there has been no change in the flux through the contour moving with the fluid. The argument can be applied to any contour moving with the fluid,

which implies that the lines of force of the magnetic field must move with the fluid. It follows that if any two points in the fluid lie on the same line of force at a time t , they will lie on a common line of force at all subsequent times. In this way it is possible to speak of the identity of a single line of force throughout an extended period of time, which is not possible usually in fluids with small magnetic Reynolds' number.

Note that the derived Eq. (3.15) does not contain the current. The system of Eqs. (3.3) and (3.15) is between v and B , and the other field quantities J and E need not be considered explicitly. The order of magnitude of the current is

$$|J| = \frac{c}{4\pi} |\nabla \times B| = \mathcal{O}\left(\frac{cB}{\ell}\right)$$

If we regard B and ℓ as given, then in the limit of infinite conductivity (11) reduces to

$$\bar{E} = -\frac{v}{c} \times \bar{B} \quad (3.17)$$

which is a fundamental relation for hydromagnetic flow. When (3.17) is satisfied (for large R^m) the lines of force move with the fluid.

It is now possible, with the use of (3.17), to discuss the order of magnitude of E and its effects, which we have so far omitted. We have

$$\frac{\partial E}{\partial t} = \mathcal{O}\left(\frac{v^2 B}{c\ell}\right) = \mathcal{O}\left(\frac{v^2}{c^2} J\right)$$

Thus $\partial E / \partial t$ is smaller than the other terms in (3.12) by $\mathcal{O}(v^2/c^2)$. The electrostatic charge density is

$$\nabla \cdot E = 4\pi \delta$$

and

$$\delta = \mathcal{O}\left(\frac{E}{\ell}\right) = \mathcal{O}\left(\frac{vB}{c\ell}\right)$$

so that the electrostatic forces on the fluid are

$$E \delta = \mathcal{O}\left(\frac{E^2}{t}\right) = \mathcal{O}\left(\frac{v^2}{c^2} \frac{B^2}{t}\right)$$

The electrostatic forces involved are a factor of v^2/c^2 smaller than the magnetic forces.

4. COLLISIONLESS PLASMA

In this case the collision time must be the greatest characteristic time of the system. We shall derive the equations governing this system by using the guiding center approximation. The basic assumptions in this approximation are as follows:

1. The radius of gyration is small compared to the radius of curvature of the magnetic field.
2. Electric field effects are small compared to those produced by magnetic fields.
3. The magnetic field may be assumed to be reasonably static in a time interval comparable in magnitude to the gyration time.

We first assume that the particle moves in a circle whose radius is given by

$$D = u_{\perp} / \Omega$$

where

$$\Omega = qB/mc \tag{4.1}$$

where u_{\perp} is the velocity of the particle perpendicular to the field. As the particle travels in its circular trajectory, we assume that the circle drifts with a small component of velocity perpendicular to the field. This drift may result from a weak electric field perpendicular to the magnetic field, from a gradient in the magnetic field perpendicular to the field, or from curvature of the lines of force of the magnetic field. The drift velocity due to a gradient in the field density across the lines of force is given by the following expression:

$$u = 1/2 \omega_{\perp} \frac{D}{\ell}$$

where ℓ is the scale of the field defined by the following equation:

$$\frac{1}{\ell} = \frac{1}{B} \frac{dB}{dn}$$

where n is the direction normal to the field. The drift follows from the variation of the radius of curvature of the trajectory as the particle moves across the gradient in B .

In vector form this drift may be written

$$u = 1/2 \frac{m\omega_{\perp}^2 c}{qB^4} \mathbf{B} \times \nabla B^2/2 \quad (4.2)$$

The motion of the particle along curved lines of force produces a centrifugal force which must be balanced by a Lorentz force from another drift velocity. If \mathcal{K} is the curvature of the lines of force, then a velocity ω_{\parallel} along the lines leads to a drift u perpendicular to B such that

$$m\omega_{\parallel}^2 \mathcal{K} = q \frac{u}{c} B$$

which when solved for the velocity of drift yields

$$u = \frac{m\omega_{\parallel}^2 \mathcal{K} c}{qB}$$

where \mathcal{K} is the curvature of the magnetic field. This curvature can be rewritten in vector form

$$\bar{u} = \frac{m\omega_{\parallel}^2 c}{qB^4} \mathbf{B} \times \{ (\bar{B} \cdot \nabla) \mathbf{B} \} \quad (4.3)$$

Finally, an electric field perpendicular to a magnetic field will give the drift velocity:

$$\mathbf{u} = c \frac{\mathbf{E} \times \mathbf{B}}{B^2} \quad (4.4)$$

The complete drift of the particles is then composed of the following three terms:

$$\begin{aligned} \mathbf{u}_\perp = c \frac{\mathbf{E} \times \mathbf{B}}{B^2} + \frac{m\omega_\perp^2 c}{2B^4} \mathbf{B} \times \nabla B^2/2 + \\ \frac{m\omega_\parallel^2 c}{qB^4} \mathbf{B} \times \{ (\mathbf{B} \cdot \nabla) \mathbf{B} \} \end{aligned} \quad (4.5)$$

Note for a definite velocity and magnetic field the electric field is determined by (4.4). If we examine (4.5) under an increasing scale of B we see that the last two terms will become smaller and smaller as the scale increases. If we assume we have a large system, then (4.5) to lowest order will contain only the first term. The principal motion, therefore, in a collisionless plasma under the condition of large scale may be thought of as the electric drift. Or conversely, the only electric fields ($\perp B$) in a collisionless plasma are those that by a Lorentz transformation give $\mathbf{E} = 0$ in the frame of reference moving with the plasma.

Now we can argue that the free motion of electrons along the lines of force means that there can be no significant electric field $\bar{\mathbf{E}}$ along $\bar{\mathbf{B}}$, i. e., E_\parallel is small compared to E_\perp . Then $\mathbf{E} \cdot \mathbf{B} \approx 0$ and if we form the cross product of \mathbf{B} with (4.4) we obtain

$$\mathbf{E} = - \frac{\bar{v}}{c} \times \mathbf{B} \quad (4.6)$$

This is the same relation that was obtained earlier for the infinitely conducting fluid. This condition is the basis for the remarkable similarity between the collisionless plasma and the conducting fluid. In the conducting fluid it arose because the electrical conductivity shorted out any electric fields in the frame of reference moving with the fluid; in the collisionless plasma it arises because $\bar{\mathbf{E}} \cdot \bar{\mathbf{B}} \approx 0$ and the particles choose to move approximately in the frame of reference in which there is no field, whatever that frame may be.

We should like to point out, however, that this is only true in the case of large scale and is not likely to occur in many laboratory fluids.

This is one element which distinguishes astrophysical models from many laboratory models.

We shall now investigate the motion along the lines of force, as a contrast to the drift velocity which is perpendicular to the lines of force. The equation of motion is due to the centrifugal force of the circular motion of the particle around the field when the field density varies along the lines of force. The acceleration of the particle along the field is given by

$$m \frac{dv_{\parallel}}{dt} = -1/2 \frac{m\omega_{\perp}^2}{B^3} \{ \mathbf{B} \cdot [(\mathbf{B} \cdot \nabla) \mathbf{B}] \} \quad (4.7)$$

The vector quantity $\mathbf{B} \cdot [(\mathbf{B} \cdot \nabla) \mathbf{B}]$ represents the angle of divergence of the lines of force where they cross the circular trajectory of the particle. We see that particles will tend to be reflected from regions of high magnetic field density.

The motion of a particle in slowly varying large-scale fields can be described by two invariants. For slow variations of the field, the magnetic moment of the particle is a constant. This is defined in terms of the perpendicular velocity component

$$\frac{m\omega_{\perp}^2}{2B} = \text{constant} \quad (4.8)$$

We also have an invariant defined in terms of the parallel component of the velocity which is obtained by integrating this quantity along a path between two reflection points on its path,

$$\oint \omega_{\parallel} ds = \text{constant} \quad (4.9)$$

Both these are adiabatic invariants and follow from the general expression

$$\oint p dq = \text{constant}$$

We ultimately are interested in the bulk motion of the fluid. We shall,

therefore, derive the Liouville equation for a time independent but spatially dependent field. We must first define the following quantities:

- θ is the angle of the trajectory relative to the field, the so-called pitch angle
- ψ is the number density of particles with pitch angle θ
- $w_{\perp} = w \sin \theta$ perpendicular component of velocity
- $w_{\parallel} = w \cos \theta$ parallel component of velocity

If we now use (4.8) we have

$$\frac{\sin^2 \theta}{B} = \frac{\sin^2 \theta_0}{B_0} = \text{constant} \quad (4.10)$$

which shows the condition for the reflection or mirroring of a particle from regions of strong field. The particle is restricted to those regions where the $\sin^2 \theta$ is less than or equal to unity.

The Liouville equation is the quantitative expression of the fact that the Lagrangian time derivative of the particle density in phase space is zero. By considering the number of particles between θ and $\theta + d\theta$ at two points $s + ds$ and $t + dt$, we arrive at the following expression for the density of the particles:

$$0 = \frac{\partial \psi}{\partial t} + \frac{\partial \psi}{\partial s} w \cos \theta + \frac{\partial \psi}{\partial \theta} \frac{w \sin \theta}{2B} \frac{dB}{ds} - \frac{\psi}{2B} w \cos \theta \frac{dB}{ds} \quad (4.11)$$

It is sometimes convenient to use the quantity φ defined

$$\varphi = \frac{\psi}{\sin \theta} = \frac{\text{number of particles}}{\text{cm}^3 \text{ solid angle}}$$

This gives an equation of the following form

$$0 = \frac{\partial \varphi}{\partial t} + \frac{\partial \varphi}{\partial s} w \cos \theta + \frac{\partial \varphi}{\partial \theta} \frac{w \sin \theta}{2B} \frac{dB}{ds} \quad (4.12)$$

By considering the characteristic solution of only the first two terms of this equation as a partial solution, we get

$$\frac{ds}{dt} = w \cos \theta \quad (4.13)$$

We see that any variation in density in the direction of the tube propagates along the tube with the component of the particle velocity along the tube. We now consider the last two terms in this equation. This gives the characteristic solution

$$\frac{\sin^2 \theta}{B} = \text{constant} \quad (4.14)$$

and

$$\varphi = \varphi \left(\frac{\sin^2 \theta}{B} \right) \quad (4.15)$$

The first condition has already been encountered, the second one tells us that an isotropic distribution (no dependence of φ on θ) will remain so in all regions that are accessible to the particles of the distribution, and the density is uniform along the line of force.

If the particle velocity distribution is anisotropic, i. e., if φ depends upon θ , then the density also depends upon B . A tendency toward small pitch angles means higher particle densities where B is large; a tendency toward large pitch angles means higher particle densities where B is small.

5. THE BOLTZMANN EQUATION FOR COLLISIONLESS PLASMA

As part of our development of the opposite extreme of the fluid approach let us consider the collisionless Boltzmann equation, i. e., $df/dt = 0$.

First, we shall define a distribution function in phase space which is a function of position, velocity, and time: $f(x_i, u_j, t)$. This distribution function will characterize our system. It represents the number of particles lying in a differential volume element in phase space at the specified values in the argument of the distribution function.

Conservation of particles along a trajectory implies the following equality:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \frac{dx_i}{dt} \frac{\partial f_i}{\partial x_i} + \frac{du_i}{dt} \frac{\partial f}{\partial u_i} = 0 \quad (5.1)$$

where d/dt is taken along the trajectory. We can write Eq. (5.1) in terms of the velocity and force:

$$\frac{\partial f}{\partial t} + u_i \frac{\partial f}{\partial x_i} + \frac{F_i}{m} \frac{\partial f}{\partial u_i} = 0 \quad (5.2)$$

We express the force F_i in terms of the forces on the particle,

$$F_i = qE_i + \epsilon_{ijk} \frac{u_j}{c} B_k \quad (5.3)$$

We shall now define a number of integrals, representing successive moments of the distribution function in velocity space. The zero moment is a scalar and represents the number density

$$N = \int d^3u f \quad (5.4)$$

The first moment is a vector and represents the momentum of the fluid.

$$Nv_i = \int d^3u u_i f \quad (5.5)$$

where v_i is the velocity of organized motion of the fluid. The second moment is a tensor and represents the pressure,

$$p_{ij} = m \int d^3u (u_j - v_j) (u_i - v_i) f \quad (5.6)$$

The third moment is a tensor of third rank which represents the heat conductivity,

$$Q_{ijk} = m \int d^3u (u_i - v_i)(u_j - v_j)(u_k - v_k) f \quad (5.7)$$

We should like to remark that the pressure tensor and the heat conductivity tensor are moments of the random velocity relative to the velocity of organized motion, in contrast to the moments of lower order.

We shall begin by taking the zero moment of (5.2); we multiply (5.2) by unity and integrate over all permissible random velocities. This zeroth moment gives conservation of mass

$$\frac{\partial N}{\partial t} + \frac{\partial}{\partial x_i} (N v_i) = 0 \quad (5.8)$$

The first moment is obtained by multiplying the equation by u_i and by integrating over all possible values of the random velocity,

$$\begin{aligned} \frac{\partial N v_i}{\partial t} + \frac{\partial}{\partial x_j} \int d^3u u_i f u_j + \frac{q}{m} E_j \int d^3u u_i \frac{\partial f}{\partial u_j} \\ + \frac{q}{mc} \epsilon_{jkl} B_l \int d^3u u_k u_i \frac{\partial f}{\partial u_j} = 0 \end{aligned} \quad (5.9)$$

The quantity ϵ_{ijk} is known as the Levi-Civita density, defined to be zero if any two of the indices ijk are equal, and either +1 or -1 otherwise, according as ijk is an even or odd permutation of 123.

The integral which comprises the second term contains the pressure tensor, and it can be removed by the following relation:

$$u_i u_j = (u_i - v_i)(u_j - v_j) + u_i v_j + u_j v_i - v_i v_j$$

Part of the third term vanishes because the distribution function vanishes at infinity faster than order of $1/u^2$. Physically this corresponds to the condition that the particles can have no momentum in the limits of infinite velocity. The only part that survives as a result of integration by parts is a number density. The integrand of the fourth term can be written

$$u_k u_i \frac{\partial f}{\partial u_j} = u_k \left[\frac{\partial}{\partial u_j} (u_i f) - \delta_{ij} f \right]$$

We see that by arguments similar to the above only the term $\delta_{ij} f$ survives to give the momentum transfer as we defined it above. Combining these results

$$\begin{aligned} \frac{\partial}{\partial t} (N v_i) + \frac{\partial}{\partial x_j} \left(\frac{p_{ij}}{m} + N v_i v_j \right) - \frac{q}{m} E_i N \\ - \frac{q}{mc} \epsilon_{ikl} B_l (v_k N) = 0 \end{aligned} \quad (5.10)$$

This equation can be simplified by using conservation of mass,

$$mN \left(\frac{\partial v_i}{\partial t} + v_j \frac{\partial v_i}{\partial x_j} \right) = - \frac{\partial p_{ij}}{\partial x_j} + qN(E_i + \epsilon_{ijk} v_j B_k)$$

The application of the above arguments to the second moment (a tensor $v_i v_j$) extracts the following equation from the original Boltzmann equation:

$$\frac{\partial p_{ij}}{\partial t} + v_k \frac{\partial p_{ij}}{\partial x_k} = - \frac{\partial Q_{kij}}{\partial x_k} - p_{ij} \frac{\partial v_k}{\partial x_k} - p_{ki} \frac{\partial v_j}{\partial x_k}$$

$$-p_{kj} \frac{\partial v_i}{\partial x_k} + \frac{q}{m} (\epsilon_{jlm} p_{il} B_m + \epsilon_{ilm} B_m p_{lj}) \quad (5.11)$$

This equation appears somewhat complicated at first sight, but is, in fact, rather simple so far as its physical effects are concerned. The second term of the right-hand side which contains the pressure tensor represents a compression term, since it has the gradient of the velocity in it. The next two terms which also contain the pressure tensor represent shear terms. The last term, which results from the magnetic fields, contains a pressure tensor. The presence of the magnetic field in this quantity shows us that the pressure anisotropy rotates with the cyclotron frequency in the presence of such a field. The above equation then tells us that any change in the pressure tensor is due to a linear superposition of four effects: changes of the heat tensor compressional changes, changes due to shear forces, and rotations in the field. We consider a system which has compression in one direction only. The only nonzero derivative of the velocity is, say, $\partial v_x / \partial x$. We ignore heat conductivity and external fields. If we consider the pressure perpendicular to the direction of compression, we obtain a linear relationship between the density and pressure,

$$\frac{dp_{yy}}{dt} = -p_{yy} \frac{\partial v_x}{\partial x}$$

Since

$$\frac{1}{\rho} \frac{\partial \rho}{\partial t} = - \frac{\partial v_x}{\partial x}$$

we have $dp_{yy}/p_{yy} = d\rho/\rho$ and $p_{yy} \propto \rho$. Physically the pressure increases in this direction because the particles are only being concentrated rather than having their speeds increased. The pressure in the direction of compression is

$$\frac{dp_{xx}}{dt} = -3p_{xx} \frac{\partial v_x}{\partial x} \quad (5.13)$$

so that

$$\frac{dp_{xx}}{p_{xx}} = 3 \frac{d\rho}{\rho}$$

whose solution gives a cubic dependence of pressure upon density

$$p_{xx} \approx \rho^3 \quad (5.14)$$

This relation is the familiar one-dimensional adiabatic law. The rapid increase of pressure results from the fact that the thermal velocity in the x direction increases proportionally to the density. Simple consideration of a particle bouncing elastically between two walls will show that the product of its velocity times the separation of the walls is a constant. This accounts for the additional two powers in the density dependence for the pressure in the direction of compression.

On the other hand, suppose we have a system in which there are only shears, with only $\partial v_y / \partial x$ nonzero. Then

$$\frac{\partial v_i}{\partial x_i} = 0$$

We shall suppose initially the only nonvanishing component of p_{ij} is

$$p_{xx} \neq 0$$

Then the equation for the pressure reduces to the following:

$$\begin{aligned} \frac{dp_{xx}}{dt} &= 0 \\ \frac{dp_{xy}}{dt} &= \frac{dp_{yx}}{dt} = -p_{xx} \frac{\partial v_y}{\partial x} \end{aligned} \quad (5.15)$$

It is seen therefore that the shear leads to off-diagonal terms in the pressure tensor. The existence of these off-diagonal terms implies a viscous medium if there are any collisions whatsoever.

We shall now consider effects on the pressure tensor resulting from a one-component magnetic field in the z direction. The following equations result:

$$\frac{dp_{xx}}{dt} = 2p_{xy} \Omega \quad \frac{dp_{xy}}{dt} = (p_{yy} - p_{xx}) \Omega$$

$$\frac{dp_{yy}}{dt} = -2p_{xy} \Omega$$

(5.16)

where

$$\Omega = \frac{2q}{mc} B$$

We see that the differential equations for the pressure tensor represent a rotation of any anisotropy at the cyclotron frequency.

6. DYNAMICS OF AN IONIZED GAS

We shall now undertake a study of the dynamical properties of an ionized gas. The analysis which is to follow has been published in greater detail.¹ As a first step we shall compute the current density from a field of particles moving in a space dependent magnetic field. We shall separate the drift motions from the cyclotron motions about the field lines and begin with these latter circular motions. We shall then designate the following parameters:

N the number of guiding centers per unit volume

Ω the cyclotron frequency (spatially dependent)

- w the velocity of the circulating particles
- D the radius of gyration of the particles ($D \ll$ scale L of the field)
- $D = w/\Omega$

Consider the net flow of particles across the interval $(y, y + \delta y)$. We displace the centers of two circles (or orbits) with radii equal to the radius of gyration computed at the guiding centers a distance δy apart. We should like to compute the current flowing through δy due to the presence of guiding centers in the overlapping area (a differential increment). An ion moving about a guiding center located in this area will contribute an average current of $\pm e \Omega / 2 \pi$, depending upon which one of the two parts of this disjointed region it lies.

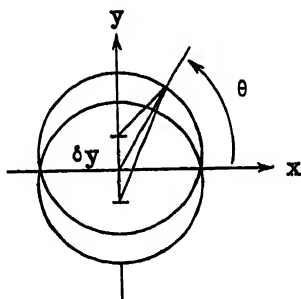


Fig. 6.1

By computing the differential area in Fig. 6.1, one can integrate over θ and arrive at the current across δy oriented with one side along the y axis. This procedure can then be extended in the z direction (perpendicular to the plane of the page) where the magnetic field has a radius of curvature of L . If we let $\delta z(x, y)$ represent the separation (measured along the magnetic field) of two neighboring surfaces orthogonal to the magnetic field, then it can be shown that the current density perpendicular to the field is

$$J_x = - \frac{c}{\delta z} \frac{\partial}{\partial y} \left(\frac{Nm w_{\perp}^2 \delta z}{2B} \right) \quad (6.1)$$

$$J_y = + \frac{c}{\delta z} \frac{a}{\delta x} \left(\frac{Nm \omega_{\perp}^2 \delta z}{2B} \right) \quad (6.2)$$

We define the pressure resulting from motions perpendicular to the field as

$$p_{\perp} = 1/2 Nm (\omega_{\perp i}^2 + \omega_{\perp e}^2) \quad (6.3)$$

i. e., the sum of electron and ion energies of circular motion.
By using the identity

$$| \bar{B} \times [(\bar{B} \cdot \nabla) \bar{B}] | = \frac{B^3}{L}$$

we can combine the component currents into the following vector equation:

$$J_c = \frac{c}{B^2} B \times \left\{ p_{\perp} - \frac{p_{\perp}}{B^2/4\pi} \nabla \left(\frac{B^2}{8\pi} \right) - \frac{p_{\perp}}{B^2} (B \cdot \nabla) B \right\} \quad (6.4)$$

This is the current arising from the circular motions of the particles around the field.

We now consider the drift velocity perpendicular to the magnetic field

$$u_d = e \frac{E \times B}{B^2} + \frac{1}{2} \frac{m \omega_{\perp}^2}{q B^4} B \times \nabla \left(\frac{B^2}{2} \right) + \frac{m \omega_{\perp}^2 c}{q B^4} B \times \{ (B \cdot \nabla) B \} \quad (6.5)$$

To obtain the current we must multiply by the number of particles and their charge. Since electrical neutrality dictates equal numbers of positive and negative particles, the first term, which is the electric drift, will give no contribution, as it is proportional to the charge of the particle. Both ions and electrons drift in the same direction. This is why a

collisionless plasma has no conductivity perpendicular to the field, i. e., because in the presence of an electric field there is no net transport of charge, only a transfer of the net bulk of the plasma. The second term will be seen to be proportional to the perpendicular pressure when it is summed over all particles. Similarly, in the third term, the sum of the parallel velocities leads to the expression for the parallel pressure. The contribution of these terms is sometimes called the "pressure current." The perpendicular pressure has two components (x and y) with equal mean-square values and can be written in the following form:

$$p_{\perp} = \frac{N}{2} m \omega_{\perp}^2 = 1/2 N m (\omega_{\perp x}^2 + \omega_{\perp y}^2) = N m \omega_{\perp x}^2 \quad (6.6)$$


and

$$p_{\parallel} = N m \omega_{\parallel}^2 \quad (6.7)$$

We can combine these quantities into a vector equation for the drift current,

$$N q u = \frac{c}{B^2} B \times \left\{ \frac{p_{\perp}}{2B^2} \nabla B^2 - \frac{p_{\parallel}}{B^2} (B \cdot \nabla) B \right\} \quad (6.8)$$

This is the current perpendicular to the field due to the drift of the particles. We shall now look at the current parallel to the field. If the particle motions are isotropic we found a uniform density distribution along the field. It is conceivable that an electric field can exist along the direction of the magnetic field. This will be true if the magnetic field separates and converges again such that two mirror points are created for the orbiting particles. In this case we may in some cases have charge separation and electric fields. The question is, "Are they important?" By looking at the force on these particles, we recall that the force exerted on the particles due to the force of the mirror was the ratio of the energy of circular motion to the characteristic scale of the field,



$$F = \frac{m \omega_{\perp}^2}{L} \quad (6.9)$$

The electric field produced by this force is the following:

$$E_{\parallel} = \frac{Nw_1^2}{cBqL} Bc = \left(\frac{D}{L}\right) \frac{wB}{c} \quad (6.10)$$

We separate out the radius of gyration so that the fields can be compared. We see that the other electric fields that might be present are due to the motions of the plasma and are of the order of $(w/c) \times B$. Since $(D/L) \ll 1$, the factor D/L means that E_{\parallel} is, in general, small compared to E_{\perp} . Finally the existence of this field E_{\parallel} depends on the existence of

anisotropy which can be maintained at a significant level only when the plasma is undergoing relatively violent motion. We shall have more to say on the anisotropy later.

We now consider the final contribution to the current. We consider a magnetic field in the z direction which is a slowly varying function of time. Perpendicular to this field we suppose that there is an electric field E_{\perp} (say along the y direction). It too is slowly varying in time. Physically, we know that the individual particle will spiral along the x direction making circles in the y, z plane as it moves. The equations of motion for the particle are:

$$\frac{d^2 x}{dt^2} = \frac{dy}{dt} \Omega \quad (6.11)$$

$$\frac{d^2 y}{dt^2} = \frac{q}{m} E - \frac{dx}{dt} \Omega \quad (6.12)$$

We divide the second equation by Ω and differentiate it with respect to time, using the first equation to eliminate dx/dt ,

$$\frac{d}{dt} \left(\frac{1}{\Omega} \frac{d^2 y}{dt^2} \right) = c \frac{d}{dt} \left(\frac{E}{B} \right) - \Omega \frac{dy}{dt}$$

We are interested in a slowly varying field such that the principal motions of the particles are circular. Then, taking a time average, the left-hand side will vanish. It follows that the mean value of the motion in the y direction due to an increase in the electric field or a decrease in the magnetic field is a migration in the y direction. The positive charges move in one direction whereas the negative ones move in the opposite. We designate this velocity as the polarization drift for this reason.

The energy that the particles receive as they move in the y direction is precisely the kinetic energy of the electric drift $c [\bar{\mathbf{E}} \times \bar{\mathbf{B}}/B^2]$ in the x direction, i. e., the electric drift kinetic energy comes from the polarization drift.

The net-polarization current can be written

$$\bar{\mathbf{J}}_p = \frac{Nmc^2}{B^2} \bar{\mathbf{B}} \times \frac{d\bar{\mathbf{u}}_d}{dt} \quad (6.14)$$

This represents the final contribution to the current. In summing up all contributions we find that a few currents cancel one another. The complete current is the following:

$$\begin{aligned} \bar{\mathbf{J}}_{\perp} = & \frac{e}{B^2} \bar{\mathbf{B}} \times \left\{ \nabla p_{\perp} + \left[\frac{p_{\parallel} + p_{\perp}}{(B^2/B\pi)} \right] \right. \\ & \left. \times \left(\frac{\bar{\mathbf{B}} \cdot \nabla}{4\pi} \right) \bar{\mathbf{B}} + Nm \frac{d\bar{\mathbf{u}}_d}{dt} \right\} \end{aligned} \quad (6.15)$$

The contributions to the current that results from the drift of the guiding centers as a result of the variation of the magnetic field density in a direction perpendicular to the field exactly cancel the contributions to the current as a result of variations of the cyclotron frequency with position.

We shall now use Maxwell's equations to obtain an expression for the field. The electric field perpendicular to the magnetic field can be found by means of the Maxwell equation,

$$\frac{\partial \bar{\mathbf{E}}}{\partial t} = c (\nabla \times \bar{\mathbf{B}})_{\perp} - 4\pi \bar{\mathbf{J}}_{\perp} \quad (6.16)$$

Note the following vector identity:

$$\frac{(\nabla \times \mathbf{B}) \times \mathbf{B}}{B} = \left[(\mathbf{B} \cdot \nabla) \mathbf{B} - \nabla \left(\frac{B^2}{2} \right) \right] \frac{1}{B} \quad (6.17)$$

Substitution for the current density into (6.16) yields the following differential equation (we have assumed that the electric field along the direction of the magnetic field is negligible):

$$\begin{aligned} \frac{\partial \mathbf{E}}{\partial t} = & \frac{c\mathbf{B}}{2(B^2/8\pi)} \times \left\{ -Nm \frac{d\bar{u}_d}{dt} - \nabla \left(p_{\perp} + \frac{B^2}{8\pi} \right) \right. \\ & \left. + \frac{(\mathbf{B} \cdot \nabla) \mathbf{B}}{4\pi} \left[1 + \frac{p_{\perp} - p_{\parallel}}{B^2/4\pi} \right] \right\} \end{aligned} \quad (6.18)$$

The right-hand side of this expression contains quantities of the order of cB/L , whereas we have shown that the variation in the electric field is small, $\mathcal{O}(v^2/c^2)$, compared to this ratio. Thus the right-hand side of the equation consists of a difference of quantities of order cB/L , all of which are very close to one another such that in combination they produce a sum smaller by $\mathcal{O}(v^2/c^2)$. We, then, can for nonrelativistic velocities ignore the left-hand side. What results then is an equation of motion from the left-hand side,

$$Nm \frac{d\bar{u}_d}{dt} = -\nabla_{\perp} \left(p_{\perp} + \frac{B^2}{8\pi} \right) + \frac{[(\mathbf{B} \cdot \nabla) \mathbf{B}]_{\perp}}{4\pi} \left[1 + \frac{p_{\perp} - p_{\parallel}}{B^2/4\pi} \right] \quad (6.19)$$

This is the equation for the electric drift. There are other contributions to the particle drift, but in the limit of small D/L they may be neglected to a first approximation. Thus we may regard \bar{u}_d as the principal bulk motion of the plasma. Now the electric drift is

$$\bar{u}_d = c \frac{\mathbf{E} \times \mathbf{B}}{B^2} \quad (6.20)$$

If we assume that

$$\vec{E} \cdot \vec{B} = 0 \quad (6.21)$$

then we obtain

$$\vec{E} = -\vec{u}_d \times \frac{\vec{B}}{c} \quad (6.22)$$

By using the second Maxwell equation

$$\frac{\partial \vec{B}}{\partial t} = -c \nabla \times \vec{E} \quad (6.23)$$

we finally obtain

$$\frac{\partial \vec{B}}{\partial t} = \nabla \times (\vec{u}_d \times \vec{B}) \quad (6.24)$$

This is the equation which relates the field to the drift velocity.

If we neglect the anisotropy $p_{\perp} - p_{\parallel}$, we see that u_d and B satisfy the same equations as in a perfectly conducting fluid, i.e., the limits of many collisions and of no collisions have led to the same results in the limit of small D/L .

Now consider the term in the equation of motion involving the anisotropy. The factor $[(\vec{B} \cdot \nabla) \vec{B}]_{\perp} / 4\pi$ represents the forces exerted by the tension $B^2/4\pi$ along the lines of force. The existence of anisotropy ($p_{\perp} \neq p_{\parallel}$) multiplies the tension by $1 + (p_{\perp} - p_{\parallel}) 4\pi/B^2$. Note then that if p_{\parallel} should exceed p_{\perp} by as much as

$$p_{\parallel} - p_{\perp} > B^2/4\pi \quad (6.25)$$

which may be very little in the presence of a weak magnetic field, the effective tension changes sign and becomes a compression. The resulting buckling in the lines of force is known as the "hose instability." When this instability sets in, there is an increase in the magnetic-field energy at the expense of the anisotropy, and isotropy tends to be preserved by the

hose instability, preventing p_{\parallel} from exceeding p_{\perp} by more than $B^2/4\pi$.

There is also an instability which occurs when $p_{\perp} \neq p_{\parallel}$, known as the "mirror instability." Briefly the mirror instability may be described as the tendency of plasma with $p_{\perp} > p_{\parallel}$ to congregate in the regions of weak field in any perturbation, because the individual particles tend to have large pitch angles and reflect from the regions of strong field. This tendency was discussed earlier with the Liouville equation. The increased plasma density in regions of weak field tends to inflate the regions of weak field, further weakening the field and trapping more plasma from the regions of stronger field, etc. The result is a growth of small scale perturbations at the expense of the anisotropy $p_{\perp} > p_{\parallel}$.

Altogether, then, we see that the presence of a weak magnetic field leads to approximate isotropy of the thermal motions in the collisionless plasma. The equation of motion then reduces to the form

$$\rho \frac{d\bar{u}_d}{dt} = -\nabla_{\perp} p_{\perp} + \frac{(\nabla \times \bar{B}) \times \bar{B}}{4\pi} \quad (6.26)$$

as for the classical fluid. The motions along the field take place freely, with pressure equalization at the speed of sound, demonstrated earlier from the Liouville equation. The magnetic field follows from Eq. (6.24), and the magnetic lines of force are permanently connected and move with the fluid velocity \bar{u}_d just as in the perfectly conducting fluid. It follows that, for a first rough approximation, the large-scale ($L \gg D$) motions of a collisionless plasma containing a weak magnetic field ($B^2/8 \ll p$) may be treated with the conventional hydromagnetic equations.

7. THE GRANULATION OF THE SUN

There is visible in the solar photosphere the convective phenomenon known as the granulation. The granulation represents the upper surface of the convective motions in the ionization zone beneath the photosphere.

Photographs of the granulation may be seen in "The Sun",^{2a} and in a recent paper by Bahng and Schwarzschild.² The convective motions represent a crude heat engine running in the temperature gradient of the general radiation field in the solar interior. It is becoming evident that the granule motions supply the energy for the extension of the solar atmosphere outward into interplanetary space, the interplanetary magnetic field, and

energy for what has been called the quiet-day "solar corpuscular radiation." It is not possible to write a quantitative theory of these effects but the basic idea is that the granule motions generate acoustical, hydromagnetic, and internal-gravity wave motions, which propagate upward through the solar atmosphere. The dissipation of these wave motions leads to the high temperature of the solar corona. The 10^6 °K temperature of the solar corona leads to supersonic expansion at several hundred kilometers per second, so that interplanetary space is filled with outward moving coronal gas. This solar wind (or solar corpuscular radiation, if one insists on a microscopic view) carries with it the weak magnetic fields enmeshed in the solar corona and forms the interplanetary medium and fields, all blowing outward from the sun at 200 to 500 km/sec.

To describe this sequence of dynamic processes leading to the solar wind, one begins with the convection beneath the photosphere.

Convection is, in general, a very difficult problem except in the most idealized forms. For the purposes of illustration, we shall assume the following simplifications:

1. The fluid is incompressible

$$\nabla \cdot \bar{v} = 0 \quad (7.1)$$

2. The temperature gradient is a constant.

3. The viscosity and thermal conductivity are constant.

We shall, then, apply a perturbation approach where the temperature and density change very little from their average values. This then gives us the following equations to work with:

$$\rho = \rho_0 + \rho_1$$

$$T = T_0 + T_1$$

$$\rho_1 = -\rho_0 \alpha T_1$$

(7.2)

$$\frac{\partial T}{\partial z} = -\beta < 0$$

(7.3)

where α is the coefficient of thermal expansion, taken to be constant. The momentum equation for the fluid is written in the form

$$\rho \frac{\partial \bar{v}}{\partial t} + \nabla p = \rho \bar{g} + \rho \nu \nabla^2 \bar{v} \quad (7.4)$$

The momentum equation for the average values is the following:

$$\nabla p_0 + \rho_0 \bar{g} = 0$$

The above equations can now be combined to give an equation for the perturbation in the fluid,

$$\rho_0 \frac{\partial \bar{v}}{\partial t} = -\nabla p_1 + \rho_0 \alpha g T_1 \bar{z} + \rho_0 \nu \nabla^2 \bar{v} \quad (7.5)$$

We throw away terms quadratic in the velocity as we are going to consider only slow motions. Following Pellew and Southwell,³ we differentiate the x component with respect to x, the y component with respect to y, and add. Using (7.1) to eliminate

$$\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y}$$

gives us

$$\left(-\frac{\partial}{\partial t} - \nu \nabla^2 \right) \frac{\partial v_z}{\partial z} = -\frac{1}{\rho_0} \nabla_{xy}^2 p_1 \quad (7.6)$$

where

$$\nabla_{xy}^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

Operating on the z component of (7.5) with ∇_{xy}^2 , on (7.6) with $\partial / \partial z$, and adding serves to eliminate p_1 , giving

$$\left(\frac{\partial}{\partial t} - \nu \nabla^2\right) \nabla^2 v_z = \alpha g \nabla_{xy}^2 T_1 \quad (7.7)$$

Operating on this equation with $[(\partial/\partial t) - \kappa \nabla^2]$ and using the conductive-heat-transfer equation

$$\frac{\partial T_1}{\partial t} - \nu_z \beta = \kappa \nabla^2 T_1 \quad (7.8)$$

we obtain

$$\left(\frac{\partial}{\partial t} - \kappa \nabla^2\right) \left(\frac{\partial}{\partial t} - \nu \nabla^2\right) \nabla^2 v_z - \alpha \beta g \nabla_{xy}^2 v_z = 0 \quad (7.9)$$

This equation may be used to treat either the onset of convection or the stationary state.

We shall now investigate the conditions for marginal instability by looking at time independent solutions. In which case the equation reduces to the following:

$$\nabla^6 v_z - \frac{\alpha \beta g}{\kappa \nu} \nabla_{xy}^2 v_z = 0 \quad (7.10)$$

Consider conditions in the layer of fluid bounded by

$$z = 0 \quad \text{and} \quad z = l$$

Since our system is homogeneous in the x and y directions, we can assume that the x and y part of our solution is sinusoidal. v_z then satisfies the

following equation

$$\nabla_{xy}^2 v_z + \frac{\zeta^2}{l^2} v_z = 0 \quad (7.11)$$

where ζ is a dimensionless constant. The equation then reduces to the following sixth-order equation:

$$\nabla^6 v_z + \frac{\zeta^2 \alpha \beta g}{l^2 \kappa \nu} v_z = 0 \quad (7.12)$$

The appropriate boundary condition is a free boundary in which case the velocity normal to both the surfaces vanishes, and shears are not permitted at the surfaces. Formally this means

$$v_z \left(\begin{array}{l} z = 0 \\ z = l \end{array} \right) = 0$$

$$\frac{\partial v_x}{\partial z} = \frac{\partial v_y}{\partial z} = 0 \quad (7.13)$$

and thus

$$\frac{\partial}{\partial x} \frac{\partial v_x}{\partial z} = \frac{\partial}{\partial y} \frac{\partial v_y}{\partial z} = 0 \quad (7.14)$$

Then since

$$\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} = -\frac{\partial v_z}{\partial z}$$

we obtain

$$\frac{\partial^2 v_z}{\partial z^2} = 0 \quad (7.15)$$

We represent the individual convective modes as follows:

$$v_z = v \sin\left(\frac{n\pi z}{l}\right) \sin k_x x \sin k_y y \quad (7.16)$$

The differential equation will accept this solution provided that

$$\zeta^2 = (k_x^2 + k_y^2) l^2 \quad (7.17)$$

If we define the Rayleigh number as

$$\underline{R} = \frac{\alpha \beta g l^4}{\kappa \nu} \quad (7.18)$$

we then obtain the following equation:

$$\underline{R} = (\zeta^2 + \kappa^2 \pi^2)^3 / \zeta^2 \quad (7.19)$$

This equation can be derived in the following manner:

$$\nabla_{xy}^2 v_z = (k_x^2 + k_y^2) v_z$$

so

$$\nabla_{xy}^6 v_z = (k_x^2 + k_y^2)^3 v_z$$

and

$$\nabla^2 v_z = (n^2 \pi^2 + \nabla_{xy}^2) v_z \quad (7.20)$$

The minimum value of \underline{R} occurs for $2, \zeta^2 = n^2 \pi^2$ yielding

$$\underline{R} = \frac{27}{4} n^4 \pi^4 \quad (7.21)$$

so that a stationary state (marginal stability) occurs in the lowest mode $n = 1$ for

$$\underline{R}_c = \frac{27}{4} \pi^4 \quad (7.22)$$

Note that the critical Rayleigh number increases as the fourth power of n , being then sixteen times larger for the second mode than for the first. Laboratory experiments by Malkus^{3a} with fixed upper and lower boundaries exhibit the vigorous onset of successively higher modes at about the expected minimum critical Rayleigh numbers for each mode. For discussion of fixed boundaries, as well as of cells with triangular and hexagonal boundaries the reader is referred to the general work by Chandrasekhar.⁴

Let us take a look at conditions in the sun, and see if \underline{R} is near the critical value. If one calculates the Rayleigh number for conditions in the photosphere, a value of about 10^{10} is obtained. So that no matter how crudely this number was obtained it is seven magnitudes greater than the critical Rayleigh number, this ensuring a state of convective motion in the photosphere. In computing this value one should be careful of the value of thermal conductivity (see reference 5) because the main mode of heat transfer is by radiative transfer processes rather than by kinetic collisional methods.

A more quantitative model of the sun's convection would require refinements in the following directions:

- (a) The actual convection motions are rapid with a large Reynolds' number, so that the nonlinear terms cannot be neglected.
- (b) The condition of incompressibility is certainly not a realistic one even though the Mach number is small because of the enormous variation of the ambient density from the bottom to the top of each convective cell.
- (c) The temperature variations across a granule are large so that β , ν , κ , and α vary significantly.
- (d) Consideration of eddy viscosity from the expected turbulence may lead to Rayleigh numbers as low as a 100.

(e) A magnetic field of 1 gauss is observed in the photosphere and, in view of the high magnetic Reynolds number, it cannot be ignored.

8. ACOUSTICAL TRANSPORT OF CORONAL ENERGY

Now consider the acoustical radiation from the convective motions represented by the granules. In terms of a general stress tensor S_{ij} , the hydrodynamic equation may be written

$$\frac{\partial}{\partial t} \rho v_i = - \frac{\partial p}{\partial x_i} - \frac{\partial S_{ij}}{\partial x_j} \quad (8.1)$$

where

$$S_{ij} = \rho v_i v_j + \delta_{ij} \frac{B^2}{8\pi} - \frac{B_i B_j}{4\pi} \quad (8.2)$$

The tensor S_{ij} may include as many stresses as we care to write down, including viscosity and gravity. If now we differentiate the equation of continuity

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_i} (\rho v_i) = 0 \quad (8.3)$$

with respect to time and use the momentum equation to eliminate $\partial/\partial t (\rho v)$, we obtain

$$\begin{aligned} & \frac{\partial^2 \rho}{\partial t^2} - a^2 \frac{\partial^2 \rho}{\partial x_i \partial x_i} \\ &= \frac{\partial^2 S_{ij}}{\partial x_i \partial x_j} + \frac{\partial^2}{\partial x_i \partial x_j} (p - a^2 \rho) \end{aligned} \quad (8.4)$$

If it is assumed that the pressure fluctuations are adiabatic, then the last term on the right-hand side vanishes. If it is assumed that the fluid motions are of sufficiently low Mach number, then the convection of acoustical waves by the fluid motion may be neglected and the left-hand side is the propagation operator for acoustical waves, giving the inhomogeneous wave equation,

$$\frac{\partial^2 \rho}{\partial t^2} - a^2 \frac{\partial^2 \rho}{\partial x_i^2} = \frac{\partial^2 S_{ij}}{\partial x_i \partial x_j} \quad (8.5)$$

To the same degree of approximation it follows that the subsonic flow v_i is relatively unaffected by the slight emission of acoustical waves so that $\partial \rho / \partial t \approx 0$ in the nonacoustical portion of the velocity field. Thus, taking the divergence of the momentum equation and using the continuity equation to eliminate $\partial(\rho v_i) / \partial x_i$, we have

$$\frac{\partial^2 S}{\partial x_i^2} = - \frac{\partial^2 p}{\partial x_i^2} \quad (8.6)$$

The inhomogeneous wave equation may be simplified to

$$\frac{\partial^2 \rho}{\partial t^2} - a^2 \frac{\partial^2 \rho}{\partial x_i^2} = - \frac{\partial^2 p}{\partial x_i^2} \quad (8.7)$$

It follows that the acoustical radiation results from the pressure fluctuations no matter how complicated and numerous may be the terms in S_{ij} . Unfortunately the Laplacian of p is one property of p that cannot be computed from the linearized equation of motion $\rho \partial v_i / \partial t = - \partial p / \partial x_i$ for incompressible flow, since $\nabla^2 p$ vanishes identically in this approximation, so it is usually convenient to write the inhomogeneous wave equation in terms of the velocity potential ϕ for the acoustical velocity field u . Then $\bar{u} = - \nabla \phi$ and

$$\frac{\partial \rho}{\partial t} = - + \rho \nabla^2 \xi \quad (8.8)$$

It follows that

$$\frac{\partial^2 \xi}{\partial t^2} - a^2 \frac{\partial^2 \xi}{\partial x_i^2} = - \frac{1}{\rho} \frac{\partial p}{\partial t} \quad (8.9)$$

so that the inhomogeneous term can be computed from the linearized equations of motion if desired. Equation (8.5) was given by Lighthill⁶ and by Kulsrud.⁷ Equation (8.6) was given by Parker.⁸ Kulsrud has pointed out that one must be careful in applying (8.5), or its equivalent, because the neglected terms may make a large contribution in the integration over retarded time.

It is interesting to work out the acoustical emission from a pulsating sphere, whose radius is given by

$$b \left[1 - \left(\frac{p_0}{\rho \omega^2 b^2} \right) \cos \omega t \right]$$

and an oscillating sphere of fixed radius b whose center occupies the position $x = 0$, $y = 0$, $z = -(2p_0/\rho \omega^2 b) \cos \omega t$. We suppose that $(p_0/\rho \omega^2 b) \ll 1$ and obtain from either (8.7) or from (8.9)

$$\xi = \frac{p_0 b}{\rho R \omega} \sin \omega t' \quad (8.10)$$

and

$$\xi = \frac{p_0 b}{\rho R \omega} \frac{\omega b}{a} \cos \theta \cos \omega t' \quad (8.11)$$

respectively, at a large radial distance R , where $t' = t - R/a$. The total

power radiated is $(2\pi p_0^2 b^2 / \rho a)$ and $(2\pi p_0^2 b^2 / \rho a) (wb/a)^{2/3}$.

The same problem can be solved by using spherical Bessel functions and boundary value methods (see, for instance, "Vibration and Sound"⁹), so that the results of the inhomogeneous wave equations can be checked. The power radiated from the oscillating sphere (a pressure dipole) is smaller than from the pulsating sphere (a pressure monopole) by the factor $(wb/a)^2 \ll 1$ because the pressure field of the oscillating sphere is dipole in character, having a $\cos \theta$ dependence. The interference between the opposite sides of the pressure fields leads to the reduction.

The radiation from the pressure monopole

$$p = p_0 \exp(-t^2/\tau^2) \exp(-r^2/b^2) \quad (8.12)$$

which might represent the pressure fluctuation in a turbulent eddy, is readily shown to be

$$\ddot{\xi} = \frac{\sqrt{\pi} p_0 b^3}{2\rho_0 a^2 \tau R} \frac{t'}{t} \exp\left(-\frac{t'^2}{t^2}\right) \quad (8.13)$$

The total energy radiated is

$$\mathcal{E} = \frac{3\pi^{5/2}}{4\sqrt{2}} \frac{p_0^6 b^6}{\rho a^5 \tau^3} \quad (8.14)$$

so that the mean power \mathcal{E}/τ is smaller than from the pulsating sphere by $(b/a\tau)^4$, or $(bw/a)^4$. The reduction is not due to interference, but to the small extent b , of the gaussian pressure field compared to the enormous extent $p_0 b/r$ of the pressure field of the pulsating sphere. In

fact, with the pulsating sphere the pressure extends so far into space that some care must be taken to avoid divergence of the retarded integral.

Altogether it is to be expected that p is of the order of $(1/2) \rho v^2$ in disordered convective motions so that with a characteristic

scale ℓ , we have $c = \ell/v$ and the radiated power is

$$P = \alpha \frac{\rho \langle v^2 \rangle^4}{a^5 \ell} \quad (8.15)$$

where α is a constant of the order of unity. The value of α suggested by the Gaussian pressure fluctuation is of the order of 10. Extensive calculations by Proudman¹⁰ suggest that α may lie in the range from 13 to 38. Kulsrud suggested that in the presence of a turbulent magnetic field, α may be as large as 10^2 , but unfortunately this increase is the direct result of working from (8.5) and treating the Reynolds and Maxwell stresses as independent so that, upon squaring and adding the resulting amplitudes, an increase was inevitable. The question remains open, then, as to whether a turbulent magnetic field will increase or decrease the net pressure fluctuations, on which the acoustical radiation depends.

With typical photospheric values of $\rho = 3 \times 10^{-7}$ gm/cm³ and $(\langle v^2 \rangle)^{1/2} = 2$ km/sec for the granules, the conservative value of $\alpha = 10$ gives a total acoustical power of $P\ell = 5 \times 10^7$ ergs/cm² sec from the granule layer. The total energy requirement of the solar chromosphere and corona was estimated by Osterbrock¹¹ to be about 2×10^7 ergs/cm³. We have more to say on the production and propagation of acoustical waves in the sequel.

9. ACOUSTICAL AND INTERNAL-GRAVITY WAVES IN THE SOLAR ATMOSPHERE

Consider the problem of the propagation of plane parallel hydrodynamic waves of small amplitude in an idealized isothermal atmosphere in the presence of a gravitational acceleration g . Orienting the coordinate system so that g is in the negative z direction and the wave vector k lies in the yz plane, the linearized hydrodynamic equations may be written

$$\frac{\partial \rho}{\partial t} + v_z \frac{d\rho_0}{dz} + \rho_0 \left(\frac{\partial v_x}{\partial x} + \frac{\partial v_z}{\partial z} \right) = 0 \quad (9.1)$$

$$\rho_0 \frac{\partial v_y}{\partial t} + \frac{\partial p_1}{\partial y} = 0 \quad (9.2)$$

$$\rho_0 \frac{\partial v_z}{\partial t} + \frac{\partial p_1}{\partial z} + \rho_1 g = 0 \quad (9.3)$$

$$\frac{\partial p_1}{\partial t} + v_z \frac{dp_0}{dz} = a^2 \left(\frac{\partial p_1}{\partial t} + v_z \frac{\partial \rho_0}{\partial z} \right) \quad (9.4)$$

for the adiabatic perturbation v_x , v_z , ρ_1 , p_1 in the unperturbed atmosphere density ρ_0 and pressure p_0 . The hydrostatic equation

$$\frac{dp_0}{dz} = -\rho_0 g \quad (9.5)$$

leads to

$$\frac{p_0(z)}{p_0(0)} = \frac{\rho_0(z)}{\rho_0(0)} = \exp \frac{-z}{H} \quad (9.6)$$

where H is the scale height kT/mg and m is mass of the individual atom or molecule in grams. If we assume that the variation of v_x , v_z , and of ρ_1/ρ_0 and p_1/ρ_0 is of the form $\exp [i(\omega t + k_x x + k_z z)]$, the dispersion relation

$$0 = \omega^4 - \omega^2 \left[a^2 (k_x^2 + k_z^2) + \frac{\gamma^2 g^2}{4a^2} \right] + g^2 (\gamma - 1) k_z^2 \quad (9.7)$$

follows, where γ is the ratio of the specific heats and $a^2 = \gamma(p_0/\rho_0)$. The substitution $iK_z = ik'_z + (1/2H)$ has been made to give real ω . The significance of the substitution is that the amplitude of the waves increase with height proportional to $1/\rho_0^{1/2}$, i. e., as $\exp (z/2H)$, as one would

expect from energy considerations. Solving for ω^2 , for a given k_x and k_z' , yields

$$\omega^2 = 1/2 \left\{ a^2 (k_x^2 + k_z'^2) + \omega_a^2 \pm [a^2 (k_x^2 + k_z'^2) + \omega_z^2] - 2 \omega_b^2 a^2 k^2 \right\} 1/2 \quad (9.8)$$

where

$$\begin{aligned} \omega_a &= \frac{\gamma g}{2a} & \omega_b &= (\gamma - 1)^{1/2} \left(\frac{g}{a} \right) \left(\frac{k_x}{k} \right) \\ k^2 &= k_x^2 + k_z^2 \end{aligned} \quad (9.9)$$

Consider the family of waves resulting when the plus sign is taken. For large wave number the dispersion relation reduces to $\omega^2 = a^2 k^2$, which is the ordinary acoustical wave. The minimum frequency occurs in the limit as $k \rightarrow 0$, yielding $\omega^2 = \omega_a^2$, showing that acoustical waves with a frequency below ω_a do not propagate in a stable isothermal atmosphere. In the solar atmosphere above the photosphere ω_a reaches a maximum in the temperature minimum where $a \approx 7.5$ km/sec. Then, with $\gamma = 5/3$ and $g = 2.7 \times 10^4$ cm/sec the limiting angular frequency is $\omega_a = 0.030/\text{sec}$, corresponding to a period of about 210 sec. This low frequency cutoff is of profound importance to the propagation of acoustical radiation from the granules upward into the solar atmosphere. The analysis of the granule motions by Bahng and Schwarzschild² suggests a mean granule life of 8.6 minutes or 516 seconds. Presumably the bulk of the acoustical radiation is produced near this characteristic period, which seems to be too long a period, by more than a factor of 2, to propagate through the temperature minimum. The radiation is largely reflected back toward the photosphere. There will undoubtedly be high frequency components for which $\omega > \omega_a$, but one expects that they represent only a small fraction of the total acoustical power. The usual Kolmogoroff

turbulence spectrum suggests that the acoustical power drops off as the 3.5 power of the period. Assuming that such a spectrum is applicable to the convective motion beneath the photosphere, less than one-tenth the total acoustical power output can pass the temperature minimum. This is not enough to supply a major portion of the chromospheric and coronal heating.

Now consider the waves resulting from the negative sign. The frequency approaches a maximum value ω_b in the limit of large k and k_x . At small wave numbers, $\omega = 2(\gamma - 1)^{1/2} a k_x / \gamma \approx a k_x$. The propagation velocity in the horizontal direction is $(\omega/k)(k_x/k) = a(k_x/k)^2$. Now $\omega_b \approx 0.03/\text{sec}$ near the photosphere and the temperature minimum, so

that the principal granule frequencies lie well below this asymptotic value, on the linear portion of the curve. Then if $k_x = k_z$ for a typical wave, the horizontal velocity is perhaps $0.5a$, or about 4 km/sec . This is comparable to the observed granule velocity. Bahng and Schwarzschild suggest that the correlation of the granule motions is not unlike a normal distribution function $\exp(-x^2/l^2)$ with $l \approx 500 \text{ km}$. The Fourier transform is thus $\exp(-k_x^2 l^2/4)$ suggesting a characteristic wave number k_x given by $k_x = 2/l = 0.4 \times 10^{-7} \text{ cm}^{-1}$, or a wavelength $\lambda_x = \pi/l = 1.5 \times 10^8 \text{ cm}$.

A mean life of $\tau = 516 \text{ sec}$ gives a characteristic velocity $\lambda_x/\tau = 3 \text{ km/sec}$, in close agreement with the dispersion relation for the internal gravity wave. It was from such considerations as these that Whitaker¹² has suggested the granules in the stable solar photosphere are largely gravity waves. They are forced oscillations, driven by the convective motions beneath, but the close approximation of their ω and k to the dispersion relation for the gravity wave suggests that much of their energy must be in a propagating mode. The suggestion is that it is internal gravity waves which transport out of the photosphere and up into the solar atmosphere the bulk of the energy which goes into heating the solar chromosphere and corona.

It should be noted that the temperature rises rapidly through the chromosphere and corona so that the low frequency limit ω_a for the propagation of acoustical waves decreases well below the characteristic granule frequency above altitudes of perhaps a few times 10^3 cm . Thus the gravity waves may excite acoustical waves in the chromosphere, as a consequence of the growing relative amplitude of the waves $|\rho_1/\rho_0| \propto \rho_0^{1/2}$ with increasing height. They may also excite the fast mode hydromagnetic wave, which was not interesting in the photosphere because of the small value of the Alfvén velocity there, but which may become important at higher altitudes because of decreasing ρ_0 . Thus, acoustical and

hydromagnetic waves may be important in the chromosphere and corona, but probably only because they are generated there by internal gravity waves from the photosphere.

Whitaker (1962)¹² has worked out coronal models based on heating by the dissipation of gravity waves alone. He points out that the dissipation, if we ignore the possibility of the formation of shocks, etc., is due principally to thermal conductivity. The kinematic viscosity of ionized hydrogen is about one-twentieth as large as thermometric conductivity. With the numerical value

$$\kappa(T) \approx 6 \times 10^{-7} T^{5/2} \text{ ergs/cm}^2/\text{sec} \quad (9.10)$$

for the thermal conductivity,¹³ he finds that gravity waves with the frequencies generated in the granules will not dissipate significantly until T reaches something of the order of 10^6 °K, whereupon they are rapidly dissipated by the steep temperature dependence of κ . In this way the model predicts, from the observed granule motions¹⁴, a temperature of the order of 10^6 °K for the solar corona. In such a model the chromosphere is heated principally by thermal conduction downward from the corona. And, of course, in any model thermal conduction downward into the chromosphere appears to be the principal energy loss to the corona.

10. THE HEATING OF THE SOLAR CORONA BY MEANS OF HYDROMAGNETIC WAVES

Hydromagnetic phenomena contain generally two characteristic velocities—the Alfvén velocity and the speed of sound. The Alfvén velocity represents the extent to which the magnetic field participates in the motions. They are given by the following formulas

$$C = \frac{B}{\sqrt{4\pi\rho}}$$

$$a = \sqrt{\gamma (kT/m)}$$

The speed of sound is then a measure of the extent to which the thermal motion participates in the wave propagations.

The photosphere is observed to have a large-scale magnetic field of about 1 gauss whereas the density is about 10^{-7} gm/cc, so that the Alfvén velocity is about 10^3 cm/sec. The temperature is in the range of about 6000°K . The speed of sound is about 10^6 cm/sec. Thus the role of the magnetic field is slight compared to the role of thermal motions. On the basis of this analysis one would tend to discount the role of hydromagnetic waves in their ability to carry energy from a granule in the solar photosphere. The time it takes a hydromagnetic wave to cross a granule

(10^5 sec) is about 100 times greater than the lifetime of the granule itself. There is one possible exception to this and this is in the region of a sun spot. The field in this region may exceed 1000 gauss and this would bring the Alfvén velocity to the same order of magnitude as the sound velocity; the region on the sun taken up by sunspots is quite small, but it is observed that the corona over regions of the sunspot is noticeably hotter. This could be attributed to hydromagnetic waves.

We shall now investigate this same question in the lower corona some 10,000 km above the photosphere. The dimension 10^4 km is small compared both to the radius of the sun and to the scale height of the general 1-gauss magnetic field so that the field can be assumed constant. The scale height for the density in the photosphere is about 100 km and so as one goes up even a small amount the Alfvén velocity will change by a drastic amount. A rough computation will illustrate this point.

The density in the upper part of the chromosphere is about 10^{-13} gm/cm³, whereas in the base of the corona it is about 10^{-15} gm/cm³. Thus the Alfvén velocity goes from 10^3 cm/sec in the photosphere to 10^6 cm/sec in the chromosphere to 10^7 cm/sec in the lower corona. The speed of sound is also increasing but nowhere as much as Alfvén speed. At the base of the corona where we have a temperature of a million degrees the speed of sound is about 10^7 cm/sec so that here the Alfvén velocity has caught up with the speed of sound. Further insight is gained by considering the magnetic field energy density $C^2\rho/2$ which is $B^2/8\pi$, and comparing it to the thermal energy density $a^2\rho$ which is NkT . As one goes up toward the corona the thermal energy is decreasing because of the decreasing number density and is, therefore, approaching the value of the magnetic energy density of the field. The temperature is increasing but not as fast as the number density is decreasing. Therefore once we reach the corona or the upper chromosphere we can no longer ignore the magnetic field as far as wave propagation is concerned. All propagating wave motions must take on a strong hydromagnetic character upon reaching the lower corona.

The appearance of spicules on the surface of the sun has led to

speculations that they may be a source of heating of the corona by means of hydromagnetic waves. Spicules are sporadic spikes of gas which leap upward from near the sun's surface and take up about a thousandth of the sun's total surface area. Conservative computations place the maximum energy they could deliver to the corona at about 10^5 ergs/cm²/sec. Since the corona consumes about this amount it would appear that spicule heating of the solar corona is a possibility; however, the distinction between corona and chromosphere is an artificial one which was brought about before the knowledge of the gas dynamics of the region had been analyzed. The fact that the two regions are coupled by thermal conduction which transfers 10^7 ergs/cm²/sec does not permit us to assume that the two regions are heated independently of one another. The spicules cannot account for this large total energy flux of 10^7 ergs/cm²/sec which the corona transfers back to the chromosphere by thermal conduction.

Of the three types of wave motion (acoustical, hydromagnetic, and internal gravity waves), the gravity waves are able to carry more energy than the rest in the photosphere. However, it is quite expected that as they rise upward they are transformed along with acoustical waves into hydromagnetic waves. The ultimate question (see discussion by Osterbrock, reference 11) is then how does a hydromagnetic wave dissipate. The viscosity is not important so that we must look to other dissipative mechanisms. Thermal conductivity may account for the dissipation. The thermal conductivity across the magnetic field is depressed somewhat by the magnetic field itself, so that if a wave propagates perpendicular to the magnetic field there may be dissipation; if the propagation acquires a component parallel to the field, then there may perhaps be enough dissipation to accomplish some of the heating. The characteristic time for dissipation is essentially the ratio of the wavelength squared to the thermometric conductivity. On the other hand, the Joule dissipation is negligible under ordinary circumstances. The electrical conductivity in an ionized hydrogen gas is about $10^7 \times T^{1.5}$ esu. Joule dissipation has an effective diffusion coefficient c^2/σ . Thus we have for the characteristic time,

$$\tau = \frac{\lambda^2 \sigma}{c} = \lambda^2 \times 10^{-5} \text{ sec} \quad (10.2)$$

We are dealing with dimensions of at least 1000 km so that we should have a time of at least 10^{11} sec; thus, it is much too long (300 years). The conductivity is decreased by the presence of a magnetic field; however, this decrease is not a result of collisions. The part of the conductivity which is responsible for dissipation might be termed the inverse resistivity

and this is solely a function of the number of collisions, which are unaffected by the presence of a magnetic field.

An electromagnetic wave propagating in a magnetic field tends to steepen its front. This can be inferred from the following physical arguments. Consider a wave or concentrated packet of magnetic energy propagating to the right. In the center of the packet the field density is greatest while it drops off on both sides of the wave. Since the sum of the gas pressure and the magnetic pressure is constant, the gas pressure, and hence the gas density, in the packet suffers an inverse effect; i. e., it decreases where the magnetic field is strongest. We see that the effective Alfvén velocity $B/(4\pi\rho)^{1/2}$ is greatest where the field is strongest, in the rising part of the wave. The rising part of the wave tends to overtake the weaker parts ahead. This effect leads to the growth of a shock front. If one examines this stability problem from second-order considerations of the equations of motion, then two conditions are required for rapid steepening,

$$\frac{\Delta B}{B} \gtrsim 1$$

$$p \lesssim \frac{B^2}{4\pi}$$

(10.3)

The motion of the wave is obtained by consideration of two conservation laws. The energy in a tube of flow of cross section A is conserved,

$$1/2 \rho v^2 CA = \text{constant}$$

(10.4)

and the magnetic flux through a tube of flow is conserved,

$$BA = \text{constant}$$

(10.5)

Thus a ratio of these two quantities gives

$$v \propto \rho^{-1/4}$$

(10.6)

The magnetic amplitude of the wave can be examined by arguments from

simple harmonic motion. The maximum variation of the magnetic field intensity represents variations in the potential energy of the system which must be equal to the kinetic energy, thus

$$\frac{1}{2} \rho v^2 \propto \frac{(\delta B)^2}{8\pi} \quad (10.7)$$

and thus

$$\delta B \propto \rho^{1/4} \quad (10.8)$$

Thus the magnetic amplitude of a wave decreases significantly as the wave propagates outward through the base of the corona. Thus, the waves cannot propagate from the region below the corona and have large amplitude in the corona. It follows that if hydromagnetic heating of the solar corona is important, the waves must be generated in the lower corona. A consideration of the energy flux needed leads to the conclusion that, if enough hydromagnetic waves are generated in the corona to heat it significantly, then ΔB will be comparable to B and they will dissipate their energy in the form of shock waves. If the corona is heated by means of magnetic waves (which may have originated in the form of gravity waves), then one can arrive at the figure of a million degrees for the corona by the equivalence of the magnetic energy and the random thermal energy. A field of 1 gauss is to be taken for the coronal field.

The temperature of the corona can, therefore, be deduced by two independent methods. If you take the period of a granule and then ask at what temperature will the thermal conductivity give significant dissipation for waves with comparable periods, you arrive at a temperature of a million degrees. Alternatively, $T \lesssim 10^6$ °K is also the condition to build up a shock wave in the field of 1 gauss. Both these conditions give a temperature of one million degrees. Thus either or both phenomena could be responsible for the observed 10^6 °K temperature.

11. THE STATICS OF A STELLAR ATMOSPHERE AT ONE MILLION DEGREES

We shall begin with a review of static models. In order to do this we must know the temperature variation in the atmosphere. For the purpose of this analysis we shall assume a temperature dependence upon

density of the following form:

$$T = T_0 \left(\frac{N}{N_0} \right)^{\alpha-1} \quad (11.1)$$

The condition of hydrostatic equilibrium is the following:

$$\frac{d}{dr} 2NkT = -Nmg(a) \left(\frac{a}{r} \right)^2 \quad (11.2)$$

The perfect gas law can be employed to relate the pressure to the number density by eliminating the temperature in the above assumed Eq. (11.2). The solution of the above equation for the pressure is, then,

$$p = p_0 \left\{ 1 - \frac{\alpha-1}{2\alpha} \frac{mga}{kT} \left(1 - \frac{a}{r} \right) \right\}^{\alpha/(\alpha-1)} \quad (11.3)$$

Determine the constant of integration p_0 at some reference point in the star where the temperature and pressure are known, the position of the point being a . The limiting form of the above expression as r becomes infinite is

$$p = p_0 \left[1 - \frac{\alpha-1}{\alpha} \frac{mga}{kT} \right]^{\alpha/(\alpha-1)} \quad (11.4)$$

If $(\alpha-1) mga / \alpha kT < 1$, then we have a finite pressure at infinity. If, on the other hand, we have

$$\frac{\alpha-1}{2\alpha} \frac{mga}{kT} > 1 \quad (11.5)$$

then at some point the pressure will vanish. This point represents the top of the atmosphere beyond which the pressure vanishes. In this case one can solve the above for the value of r corresponding to the top of the atmosphere. The largest value α can have is $5/3$ which is the value of α for an adiabatic system consisting of a monatomic gas. The left-hand

side of the above inequality represents the ratio of the gravitational potential energy to the random thermal energy. Therefore, if the atmosphere is too hot, then the quotient is small and the atmosphere will extend to infinity. On the other hand, if the temperature is low the atmosphere will be finite.

We shall now make some estimates for the sun. To be as conservative as possible take α equal to $5/3$, the largest possible value. If we put in values for the base of the corona, we find that conditions are well satisfied for the atmosphere being limited to a finite extent. However, it is obvious that the lower corona is not an adiabatic, but rather more nearly an isothermal system. Computing $(\alpha - 1) m g a / \alpha k T$ at larger distances from the sun, always with the conservative assumption that $\alpha = 5/3$ immediately beyond the point of computation, leads to decreasing values which reach 1.0 at the limits of observation (see reference 14 for estimates of T) at $22 R_{\odot}$. Since, in fact, we expect from the high thermal conductivity that α is more nearly equal to 1.1 or 1.2 than to $5/3$ (see later discussion), $(\alpha - 1) m g a / \alpha k T$ evidently falls well below 1.0. We conclude that the solar corona is too hot to be static.

We shall now discuss a static model due to Chapman, who was the first to suggest that the corona extends far out into space, at least as far as Earth. Though the accuracy of this model can be questioned, it serves the useful purpose of pointing out that the corona can be quite extensive because of thermal conductivity. Chapman assumed that after a few solar radii one may neglect energy loss to the corona by radiation. The temperature of the corona is then determined solely by the heat flow equation in the steady state

$$\nabla \cdot (K \nabla T) = 0 \quad (11.6)$$

where

$$K \propto T^{5/2}$$

For the simple case of spherical symmetry, we have

$$\frac{1}{r^2} \frac{\partial}{\partial r} (r^2 T^{5/2} \frac{\partial T}{\partial r}) = 0 \quad (11.7)$$

Since the thermal conductivity is independent of the density, we are able to solve the above equation independently of the barometric law. The solution is the following:

$$T \propto \frac{1}{r^{2/7}}$$

Assuming a temperature of two million degrees for the corona, the temperature at the earth's orbit (200 solar radii) is 400,000 degrees. This result is an indication of how far out the sun's corona extends. We will now integrate the barometric equation. We will do this quite generally by assuming the following spatial dependence for the temperature.

$$T = T_0 \left(\frac{a}{r} \right)^s \quad (11.8)$$

The above equation is applicable to the Chapman model if we assign the value of $2/7$ to s ,

$$\frac{dp}{dr} = 2kT \frac{dN}{dr} + 2Nk \frac{dT}{dr} = - \frac{Nmg a^2}{r^2} \quad (11.9)$$

from which we find the following barometric formula:

$$p = p_0 \exp \left[- \frac{mga}{kT_0(1-s)} \left(1 - \left(\frac{a}{r} \right)^{1-s} \right) \right] \quad (11.10)$$

Comparing the pressure at infinity to that at the base of the corona we find a ratio of $1/10,000$ which is much too large a ratio to be consistent with observation. The estimated pressure at the base of the corona is about 1 dyne/cm^2 whereas that for interstellar space at $r =$ is about $\leq 10^{-12} \text{ dyne/cm}^2$. The interstellar medium could not possibly push in hard enough to contain the enormous outward pressure $p(\infty)$ exerted by the solar corona. One sees that the main criticism of this model is that it does not permit the temperature to drop off sufficiently rapidly. This model brings out the important point that thermal conductivity is responsible for extending the solar corona out far into space. If one attempts to solve for the density, it diverges with the radius; in fact, the mass, i.e., the integrated density, at a distance of 100 au becomes 300 solar masses. Therefore, one has to postulate an inward pressure which contains the atmosphere and does not allow the pressure to go to zero. Such a postulate is not a realistic one, and, therefore, we shall undertake the study of a dynamic, rather than static, atmosphere.

12. HYDRODYNAMIC EXPANSION OF THE SOLAR CORONA

Consider the hydrodynamics of steady radial gas flow in an idealized corona with spherical symmetry about the sun.¹⁵ Conservation of mass requires that the number density $N(r)$ be related to the flow velocity $v(r)$ by the equation

$$N(r)v(r)r^2 = N_0 V_0 a^2 \quad (12.1)$$

where N_0 and V_0 refer to the values at the reference level $r = a$. We shall choose a to lie in the lower corona. The momentum equation is

$$Nm v \frac{dv}{dr} = - \frac{d}{dr} (2NkT) - \frac{GM_\odot mN}{r^2} \quad (12.2)$$

where m is the mass of a hydrogen atom. The hydrostatic pressure of the ionized coronal gas is $2NkT$. Instead of an energy equation we employ the artifice of the polytrope relation $T = T_0 (N/N_0)^{\alpha-1}$ where $\alpha-1$ is a measure of the extent to which the gas cools upon expansion during its outward motion from the sun. Observations suggest that α may be of the order of 1.0 to 1.2, though this is a very rough estimate. It can be shown that thermal conductivity is sufficient to maintain α near 1. Adiabatic conditions would be represented by $\alpha = 5/3$. With the polytrope law the momentum equation can be integrated to give Bernoulli's law, which may be written

$$\begin{aligned} V^2(\xi) + \frac{2\alpha}{\alpha-1} \frac{V_0}{V\xi^2}^{\alpha-1} - \frac{\Lambda}{\xi} &= V_0^2 + \frac{2\alpha}{\alpha-1} - \Lambda \\ &\equiv V_\infty^2 \end{aligned} \quad (12.3)$$

upon elimination of N with the equation of mass conservation. We have put $\xi = r/a$ and $V^2 = 1/2 Mv^2/kT_0$ so that V represents the gas velocity

in units of the thermal velocity at the reference level. The symbol Λ is $GM_{\odot}M/akT_0$ and represents the gravitational potential energy at the reference level in units of the thermal energy.

Now when $V_{\infty} > 0$, Eq. (12.3) has the two asymptotic branches

$$V \sim V_{\infty} \quad (12.6)$$

$$V \sim \frac{V_0}{\xi^2} \left(\frac{2\alpha}{(\alpha-1)V_{\infty}^2} \right)^{1/2} \quad (12.7)$$

at large ξ . When $\alpha < 5/3$ there is the upper asymptotic branch

$$V \sim \frac{\Lambda^{1/2}}{\xi^{1/2}} \quad (12.8)$$

at small ξ , and if $\alpha < 3/2$ there is the lower branch,

$$V \sim V_0 \frac{2\alpha}{(\alpha-1)\Lambda} \frac{1}{\xi} \frac{1}{\Lambda} \frac{(3-2\alpha)}{(\alpha-1)} \quad (12.9)$$

at small ξ . It is readily shown with the aid of (12.1), the branch represented by (12.5) leads to nonvanishing density and pressure at infinity. Thus it does not meet the boundary condition that $p \rightarrow 0$ as $r \rightarrow \infty$. Chamberlain¹⁶ has discussed the special case that the solar corona has precisely the temperature which causes V_{∞} to be identically zero, but the requirement seems artificial. Thus with $V_{\infty} > 0$ we require that the solution lie on the upper branch, given by (12.4), for which N goes to zero like $1/r^2$ as $r \rightarrow \infty$. On the other hand, it is observed that the quiet-day solar corona is relatively steady near the sun with V becoming small as we approach the base of the corona. Hence the lower branch (12.7) at small r is required. The crossover between the lower branch (12.7) and the upper branch (12.4) occurs at the "critical point" (V_c, ξ_c) . To investigate the crossover, differentiate (12.3), obtaining

$$\frac{dV}{d\xi} \left(V - \frac{\alpha V_o^{\alpha-1}}{V^{\alpha-1} \xi^{2(\alpha-1)}} \right) = \frac{2\alpha V_o^{\alpha-1}}{V^{\alpha-1} \xi^{2\alpha-1}} - \frac{\Lambda}{2\xi^2} \quad (12.8)$$

On the one hand, since the expansion velocity V_o at the reference level in the lower corona is smaller than α , the coefficient of $dV/d\xi$ is clearly negative at $\xi = 1$. The right-hand side of the equation is also negative if $\Lambda > 4\alpha$, giving $dV/d\xi$ positive. The velocity is increasing outward on the lower branch at small ξ . On the other hand, with $\alpha > 1$, both the coefficient of $dV/d\xi$ and the right-hand side are positive on the upper branch as $\xi \rightarrow \infty$. The two quantities must both pass through zero somewhere between large and small ξ . The critical point (V_c, ξ_c) is the point at which they pass through zero simultaneously; hence

$$V_c^{\alpha+1} \xi_c^{2\alpha-2} = \alpha V_o^{\alpha-1}$$

$$V_c^{\alpha-1} \xi_c^{2\alpha-3} = 4\alpha V_o^{\alpha-1} / \Lambda \quad (12.9)$$

Thus

$$V_c^2 \xi_c = \frac{\Lambda}{4} \quad (12.10)$$

and

$$V_c = \left[\frac{4\alpha V_o^{\alpha-1}}{\Lambda} \right]^{2(\alpha-1)/(5-3\alpha)} \quad (12.11)$$

Expanding $V(\xi)$ about (V_c, ξ_c) it is readily shown that (12.4) and (12.7) are connected by the solution which passes straight across (V_c, ξ_c) , for which V_0 is given by substitution of (12.10) into (12.3). The solutions $V(\xi)$ beginning with smaller V_0 reach a maximum, where the right-hand side of (12.8) vanishes, at a value of $V < V_c$, and decline at large ξ on the lower branch (12.5). Solutions beginning with larger V_0 than for the solution crossing the critical point turn around, where the coefficient of $dV/d\xi$ vanishes, at a value of $\xi < \xi_c$ and return toward $\xi = 0$. Such solutions are without physical significance.

Suppose, then, that the expansion follows the lower branch (12.5) at large ξ . The density and pressure are

$$N \sim N_0 \left(\frac{(\alpha - 1)V_\infty^2}{2\alpha} \right)^{1/(\alpha - 1)}$$

$$p \sim p_0 \left(\frac{(\alpha - 1)V_\infty^2}{2\alpha} \right)^{2/(\alpha - 1)}$$

in the limit as $\xi \rightarrow \infty$. This pressure is only a few factors of ten less than the pressure p_0 at the base of the corona. There is no known interstellar force capable of pushing in on the coronal gas to maintain this asymptotic value of p . In the absence of an inward pressure, we should imagine that the outward flow will accelerate. But this in no way relieves the difficulty, since, all other things being constant, V_∞ , and hence p at infinity, increase with increasing V_0 . Thus the acceleration must continue until V_0 reaches the value sending $V(\xi)$ through the critical point (V_c, ξ_c) and the velocity proceeds to infinity along the upper branch (12.4). On the upper branch, p falls to zero as $1/\xi^{2\alpha}$ as ξ becomes large. Thus it is the upper branch which is in equilibrium with the interstellar vacuum, and any slower expansion is expected to accelerate to the upper branch in the absence of interstellar pressure. The corona expands to form an outward flow of gas from the sun. The velocity of the gas—the "solar wind"—reaches supersonic velocities at large distances from the sun.

Clauser¹⁷ has pointed out that the expansion of the corona described by the above equations is exactly analogous to the expansion of a gas

through a Laval nozzle into a vacuum, with the gravitational field playing the role of the nozzle. This is readily demonstrated by noting that the equations for the flow through a nozzle with cross section $A(s)$ can be written

$$V^2 + \frac{2\alpha}{\alpha-1} \left(\frac{V_o A_o}{VA} \right)^{\alpha-1} = V_o^2 + \frac{2\alpha}{\alpha-1} \quad (12.12)$$

By differentiating this expression, we have

$$\begin{aligned} \frac{dV}{ds} \left[V - \frac{\alpha V_o^{\alpha-1}}{V^\alpha} \frac{A_o^{\alpha-1}}{A^{\alpha-1}} \right] \\ = \alpha \left(\frac{V_o}{V} \right)^{\alpha-1} \frac{A_o^{\alpha-1}}{A^\alpha} \frac{dA}{ds} \end{aligned} \quad (12.13)$$

In a nozzle, V starts with a small value V_o and increases as the cross section A diminishes ($dA/ds < 0$) into the throat, beyond which $dA/ds > 0$. It is readily shown that, if the nozzle exhausts into a vacuum, the gas follows the solution across the critical point, where $dA/ds = 0$, $A = A_{\min}$, and passes from the lower branch to the upper branch of $V(s)$, giving supersonic velocity in the widening portion of the nozzle beyond the throat. The exact analogy between the nozzle and the expansion of the solar corona obtains upon establishing some convenient correspondence $\xi = f(s)$ between s and ξ , such that $V(\xi) = V(s)$, and adjusting the right-hand side of (12.13) to be equal to the right-hand side (12.8). Then to duplicate a solution $V(\xi)$ we put

$$\left(\frac{A_o}{A} \right)^\alpha \frac{d}{ds} \left(\frac{A}{A_o} \right) = \frac{2}{f^{2\alpha-1}(s)} - \frac{\lambda}{2\alpha f^2(s)} \frac{V^{\alpha-1}(s)}{V_o^{\alpha-1}} \quad (12.14)$$

The integral of this equation is obtained by solving (12.12) for A/A_o .

$$\frac{A(s)}{A_0} = \frac{V_0}{V(s)} \frac{1}{[1 + \{(\alpha - 1)/2\alpha\} \{V_0^2 - V^2(s)\}]^{1/(\alpha - 1)}} \quad (12.15)$$

It would be interesting to try to duplicate the hydrodynamics of coronal expansion in the laboratory with such a nozzle.

Now let us collect together the conditions under which the corona will expand into the interstellar vacuum with supersonic velocity, with the gravitational field acting as a Laval nozzle. First of all we must require that $V_\infty > 0$, in order that the supersonic branch of (12.3) exist. Assuming that V_0 is small, this is equivalent to $\Lambda < 2\alpha/(\alpha - 1)$. If it should happen that $\Lambda > 2\alpha/(\alpha - 1)$, it is readily shown that the corona becomes a static polytropic atmosphere with a top ($N \rightarrow 0$) at some finite radial distance from the sun. Second, we require that with V_0 small ($V_0 < \alpha$), the velocity shall increase outward from $\xi = 1$. This is equivalent to the statement that the right-hand side of (12.8) must be negative at $\xi = 1$, giving the restriction $\Lambda > 4\alpha$. The requirements are, then, that the gravitational field, represented by Λ , must be strong enough to simulate the throat of a nozzle, but not so strong, of course, as to close off the flow entirely ($A_{\min} \rightarrow 0$). Thus the corona will expand supersonically if

$$4\alpha < \Lambda < \frac{2\alpha}{\alpha - 1} \quad (12.16)$$

Numerically we have $\Lambda \approx 16 \times 10^6/T$. From observations of the corona near the sun one concludes that the effective value of α is probably somewhere between 1.0 and 1.2. Thus, on the basis of this simple polytropic representation, any coronal temperature between about 1 and 4 times

10^6 K leads to supersonic expansion. Observations of the coronal temperature throughout the sunspot cycle suggest that (12.16) may always be satisfied in the solar corona. Certainly the observations of comet

tails,¹⁸ of the quiet-day aurorae, of polar geomagnetic agitation, etc., suggest that the solar wind is present throughout the entire solar cycle. Recent plasma observations from the Soviet Lunik space vehicles and from Explorers X and XII suggest that on quiet days the strength of the solar wind at the orbit of Earth is from 1 to 5×10^8 protons/cm² sec (with an equal number of electrons, of course), presumably some 10 protons/cm³ at a few hundred km/sec. The measurements so far have been more qualitative than quantitative and may be taken only to demonstrate the

existence of a continual solar wind of the above magnitude without giving precise values of the density and velocity. The difficulty has been that the Soviet plasma detectors required large corrections for photoelectric emission when looking toward the sun, and the U.S. vehicles did not get clear of the enormously broad transition region between the geomagnetic field and the interplanetary solar wind. Forthcoming space experiments in both countries are expected soon to give quantitative data.

As a final comment on the steady expansion of the solar corona, let it be emphasized that the actual corona may consist of such irregularities as hot and cold filaments, perhaps retaining their identity far into space. There may be large-scale variations of coronal temperature around the sun, and there may be many small-scale irregularities appearing in the solar wind in interplanetary space, as the result of small anisotropies in the thermal motions, shears, etc. Thus the present calculations must be viewed as a large-scale time-averaged view of the whole. It may be hoped that future observations in space will fill in the many interesting details that have been omitted in the present gross dynamical picture.

The energy involved in the quiet-day solar wind turns out to lie somewhere between 10^{27} and 10^{28} ergs/sec and the solar mass loss near 10^{12} gm/sec. The mass loss amounts to about 10^{-4} M. in the 5×10^9 year life of the sun, and may be regarded as negligible. The energy expenditure is readily accounted for by thermal conduction in a temperature gradient of the order of 10°K/km at the base of the corona where $T \approx 2 \times 10^6$ K. (These numbers correspond to a scale height of 10^5 km and $\alpha \approx 1.1$.) The total energy flux from the sun across a surface of radius r is

$$F(r) = 4\pi r^2 \left\{ N(r)v(r) [3kT + 2kT + 1/2 mv^2 - G \frac{M_\odot M}{r}] - \kappa(T) \frac{dT}{dr} \right\} \quad (12.17)$$

where $3kT$ represents the convection of thermal energy across the surface r , $2kT$ represents the work done by the pressure on the gas ahead, $1/2 mv^2$ is the kinetic energy of the bulk motion, $GM_\odot m/r$ is the gravitational potential energy, and the last term represents thermal conduction. Conservation of energy requires that $F(r)$ be independent of r . Very far from the sun we would expect that all the terms are negligible except $1/2 mv^2$.

Very near the sun $1/2 mv^2$ must be negligible, so that equating the resulting two expressions for $F(r)$ we obtain

$$N_0 V_0 \left[\frac{1}{2} m v_\infty^2 + \frac{G M_\odot M}{a} - 5kT_0 \right] = - \kappa_0 \left(\frac{dT}{dr} \right)_0 \quad (12.18)$$

where v_∞ is the solar wind velocity of $r = \infty$, and the subscript zero denotes the values at the reference level $r = a$. With $v_\infty = 300$ km/sec, $a = 10^6$ km, and $T_0 = 2 \times 10^6$ °K, the three terms on the left-hand side are $(0.8 + 2.2 - 1.4) \times 10^{-9} = 1.6 \times 10^{-9}$ erg/proton. With $NV = 2 \times 10^8$ protons/cm² sec at the orbit of Earth, we have $N_0 V_0 = 4.5 \times 10^{12}$ protons/cm² sec, so that with these numbers the total left-hand side is 7×10^3 ergs/cm² sec, representing the energy flux at the reference level. The thermal conductivity at 1×10^6 °K is about 6×10^8 ergs/cm² sec per °K/cm, and, at 2×10^6 °K, it is about 6 times as much. Thus a thermal gradient of 1 °K per km at $r = a$ supplies 6×10^3 ergs/cm² sec if $T = 1 \times 10^6$ °K, and 4×10^4 ergs/cm² sec if $T = 2 \times 10^6$ °K. The lower temperature seems to be adequate for an energy supply based entirely on thermal conduction. Thus it does not appear necessary that heat be supplied to the expanding corona above its base by any means other than thermal conduction.

13. STELLAR WINDS AND SOLAR PHENOMENA

The solar wind is due to the steady expansion of the corona. The work of Bridge and Rossi indicates that there is a proton density of about 10 protons/cc and that their average velocity is about 300 km/sec. It was once believed that the only protons that could escape from the sun to create the solar wind were those in the upper part of the Maxwellian tail whose velocities exceeded the escape velocity. However, as we have seen, one can expect the large observed fluxes on the basis of the supersonic expansion of the corona in the solar gravitational field.

It is interesting to speculate upon the possibilities of stellar winds from other stars. The sun is a population II, main sequence star with a surface temperature of approximately 6000°K. The solar wind from the corona is due, evidently, to the existence of a convective photosphere. A

theoretical investigation would indicate that such convective motion could exist in stars with temperatures less than about 8000°K . Thus all stars below middle class F stars should have this convective zone and thus produce a solar wind. This figure of 8000°K represents the value for which a great deal of collisional ionization can take place in hydrogen; thus cooling the gas and cutting down its adiabatic gradient so that the main mode of heat transfer becomes convection. Above 8000°K the ionization of hydrogen becomes complete so there are no atoms about that can ionize. Convection requires that we have un-ionized matter, which becomes ionized with increasing temperature, to raise the effective specific heat of the medium through the absorption of energy by collisional ionization. A star with a high temperature then does not permit this passage from ionized to un-ionized matter to occur.

From an observational point of view, we can examine all stars that exhibit some sort of mechanical or turbulent motion. As we have seen this mechanical motion then leads to acoustical, hydromagnetic, or gravity waves, which then in turn lead to the existence of a corona and solar wind. The higher luminosity stars above $10,000^{\circ}\text{K}$ show wide lines in their spectrum which are inconsistent with the temperature observed from the ratio of line intensities. Thus some form of turbulent motion is suggested to account for these line widths. It is not unreasonable to speculate, then, that there may be considerable wave generation and coronal heating in stars with photospheric temperatures above

10^4°K . We would expect, tentatively, that such stars have vigorous stellar winds. The next class of stars under suspicion are the Wolf Rayet and class B emission stars which are observed to have extended envelopes with rising and falling columns of gas. This massive and violent convective motion must lead to vigorous atmospheric and coronal heating, suggesting strong stellar winds. The magnetic stars seem to have conditions favorable for the propagation of hydromagnetic waves, especially since many are variable, presumably containing a pulsating magnetic field, leading to a turbulent and heated atmosphere, and stellar winds. The close double stars with their mutual periodic perturbations of each other's atmosphere would also represent a group of stars for which there should be a hot corona, and stellar winds.

The dwarf stars are observed to have no turbulence. However the sun from a distance of 1 parsec would be observed to have no turbulence either, so that the observations can not be taken as indicating that there exists no corona or stellar wind here.

In general, it is difficult to observe the solar wind directly. The observation of comet tails indicates that they are displaced from their trajectories away from the sun, and so this is the only visible evidence

for the existence of the solar wind.¹⁸ Certain giants which are cool red stars at about 3500°K have by direct observation been credited with a visible mass loss. In particular the work of Deutsch¹⁹ in his observations

of α^1 Hercules and its companion star α^2 which passes very close to it. A displacement of absorption lines in the violet occurs when α^2 passes behind α^1 , indicating velocities of about 10 km/sec. From this Deutsch estimates a mass loss of about 3×10^{-8} solar mass per year. Thus in a 3×10^7 years this star would lose 1 entire solar mass. On an astronomical time scale this is quite large. This estimate for the mass loss is based upon a radiative model which, though it can be criticized still, is the most plausible model to be suggested.

The only alternative theory to coronal expansion attempting to explain this mass loss is a catapult theory in which it is suggested that some mechanism is responsible for a sudden emission of material from the surface of the star and then radiation pressure supplies the additional energy to keep it going. This theory is motivated by the observations of flares and eruptions which occur frequently on the surface of the sun; but is without theoretical basis so far. There remains in the catapult theory to write down a fundamental Newtonian equation of motion. This leaves only the expanding atmosphere theory as a tool for theoretical research in investigating the possibility of stellar winds in other stars.

We shall now make some remarks concerning the numerics of this analysis of an expanding atmosphere. We wrote down a conservation law which related the energies at some reference level at the surface of the photosphere to the energy at infinity. Initially most of the energy was in the thermal energy of the medium and the energy at infinity was in the organized motion of the fluid. Thus we should have that

$$v_{\infty} \gtrsim \sqrt{T_0} \quad (13.1)$$

Since the velocity was cast in terms of the thermal velocity and the model has shown us that the atmosphere is essentially isothermal, we expect that the velocity (in this system of units) at infinity would be of the order of unity.

We also saw that the random thermal energy in the solar corona was some fraction of the gravitational energy, the fraction depending on our assumption about the value of the polytropic index of the gas. Generally we could say that this fraction would be about 0.1. So that we may write down the following:

$$kT \approx 0.1 \frac{GM_s M}{R_s} \quad (13.2)$$

If kT should momentarily exceed this value by very much, the enormous coronal expansion that results would soon reduce T again to this general level. On the other hand, it does not take much coronal heating to maintain T at a little below this order of magnitude. Altogether we expect the relation to apply to any star with significant coronal heating.

The ratio of the mass to the radius of a star is pretty much the same for most stars of the main sequence. Combining (13.1) and (13.2) leads to the conclusion that stellar winds blow outward from stars with velocities which are some sizeable fraction of the gravitational escape velocity from the surface of the star. This velocity is of the order of about 500 km/sec for most stars on the main sequence.

Applied to the sun, Eq. (13.2) reveals the thermostatic nature of the solar corona which when calculated gives a temperature of the order of

$10^6 \text{ }^\circ\text{K}$. This represents the third independent method by which the temperature of the sun's corona is predicted to be the same value. Three independent theoretical methods based on the propagation of internal gravity waves, hydromagnetic waves, and the existence of the solar wind, combined with observations of the 1-gauss magnetic field of the sun and the periods and wavelengths of the granules all lead to the same

$10^6 \text{ }^\circ\text{K}$ for the solar corona. All these estimates seem to occur independently of one another, leading one to speculate whether or not they are coincidental. Moreover the interesting question of whether or not other stars with convective photospheres would also yield these coincidences is one that is still open.

Interplanetary Conditions

In view of the knowledge of our previous developments we shall now speculate about the physics of interplanetary space. Of particular interest is the consequence of a solar wind to an interplanetary medium and magnetic field. We know that the slightest magnetic field will entrain an isolated hydrogen ion. Thus we could not have any stationary protons in interplanetary space unless we are willing to postulate the absence of a magnetic field in the solar wind. Neutral particles, on the other hand, are unaffected by the magnetic field. Their ionizational collision cross section is very low and requires high velocity electrons in order to separate the bound electron from its proton nucleus. Only where the solar wind

temperature is $10^4 \text{ }^\circ\text{K}$ will electrons have sufficient energy to ionize. However, the charge exchange cross section for proton-hydrogen collision is very high. The charge exchange reaction rate is fairly constant over the range of interest (from 100 to 40 kv) so that we could make an estimate of the lifetime of a neutral atom. Taking a density of about 10 particles/cc we have a lifetime for a neutral atom of about 10 days. The transit time from the sun to the earth is about 5 days. Thus a high velocity proton

passes an atom, strips the electron from it, and exchanges a little kinetic energy. The newly formed atom now has, in effect, the former proton's velocity; on the other hand, the newly formed proton is accelerated by the magnetic field. In view of this analysis one can not postulate a static background of interplanetary matter. Furthermore, the densities are so low that recombination is not significant and we can say that the state of ionization of the solar wind at the earth's orbit will be similar to that of the solar corona.

It is of interest to consider collision times of the solar wind ions, compared with the transit time from the sun to, say, the orbit of Earth. At 300 km/sec the Sun-Earth transit time is about 5×10^5 sec, or 5 days. The proton-proton collision time is $t = (w^3/N) \times 10^{-13}$ sec, where w is the proton thermal velocity. Since we have no present idea of the temperature of the solar wind at the orbit of Earth, consider the two cases $T = 10^3$ °K and $T = 10^6$ °K. Then, for $N = 10/\text{cm}^3$, we have $t \approx 2 \times 10^3$ sec and 4×10^7 sec, respectively. It turns out that if the solar wind is

as hot as 5×10^4 °K, the collision time is long compared to the transit time, and the solar wind may be regarded as collisionless in interplanetary space. This immediately opens up the possibilities of anisotropies in the thermal motions, which may lead to the mirror and hose instabilities, etc., and hence to small-scale disordering in the interplanetary magnetic fields. The motion of the solar wind still satisfies the hydrodynamic equations, of course, as noted earlier.

Considering the possibility of turbulence in the solar wind, note that the Reynolds number will be low for high temperatures, indicating nonturbulent flow, whereas for low temperature the Reynolds number is high. It is difficult to decide just when turbulence might appear because of the possibility that the corona is composed of hot and cold filaments.

The high electrical conductivity of the gas leads to a high magnetic Reynolds number at any temperature. For 10^4 °K, we have $\sigma \approx 10^{13}$ esu, so that

$$R^m = \frac{4\pi\sigma}{c} = \frac{10^{13+7+13}}{10^{21}} = 10^{12} \gg 1.$$

The magnetic Reynolds number is approximately the same for the corona. At the corona we can safely say that the expansion which creates the solar wind carries with it all the lines of force of the general 1-gauss solar field which passes through the expanding portions of the corona. In order to discuss the field in interplanetary space we must first investigate its properties on the sun's surface. We observe the field on the sun by means

of the Zeeman effect in the photosphere (which indicates a field of about 1 gauss). This method, however, does not resolve the field and can give only an average field over a large area. If you look at the sun's surface during an eclipse you observe streamers which are essentially radial. It is believed that these map out the magnetic lines of force. One concludes that the field is principally radial, except perhaps in the vicinity of active regions, where loops are often observed. The field is observed to be uniform about the poles down to a latitude of about 55° . At lower latitudes the field has no regular or fixed pattern. The observed field is everywhere of the order of 1 gauss intensity, except for regions with sunspots which will have a much larger field intensity.

In the corona we have a balance between the magnetic field energy and the random thermal motion,

$$NkT = \frac{B^2}{4\pi}$$

If the thermal energy were less than the magnetic pressure, then the field would close over the equatorial region, connecting the polar regions in dipole fashion. As it now stands the particles carry out the lines of force so that they extend into space instead of closing at the sun. The magnetic pressure drops off more rapidly with distance from the sun than the dynamic pressure of the solar wind, so that to a first approximation the magnetic stresses far from the sun can be neglected. It follows that the gas flows radially from the sun.

If we extrapolate the field of 1 gauss from the surface of the sun radially outward to the earth's orbit, we obtain a value of about 2×10^{-5} gauss. The sun rotates so that the lines of force take the form of an Archimedian spiral. Rough analytical estimates place the slope of the lines of force at the earth's orbit at about 45° relative to the radial direction. The gas flows radially, of course.

In 1958 it was observed that the magnetic field at one pole reversed itself, and, then, a few months later it was followed by the other pole reversing itself. We have only been observing the direction of the magnetic field since 1953 so that we cannot say if this reversal is coincident with the sunspot 11-year cycle.²⁰

Expansion of the Active Corona

So far we have discussed the steady quiet-day corona, this corona consisted of an expanding gas starting from a subsonic speed and accelerating in the gravitational field to supersonic speed. We shall now discuss the phenomena of an active corona, in particular, the state of the corona

following a large flare. The largest flares have a lifetime of around 30 min to 1 hr. Flares usually rise rapidly to a maximum intensity and remain there for perhaps 15 min and then slowly fade away. The total energy expenditure can run as high as 10^{32} ergs for the largest flares which, when compared to the total luminosity of the sun of 4×10^{33} ergs/sec, is quite a large value. The larger flares are as broad as 2×10^4 km. They are associated with sunspot groups. Within a day or two after the observation of a flare, particle clouds from the sun arrive at the orbit of the earth, indicating a velocity of 1 to 2 thousand km/sec. The particle density of this gas has been estimated as high as 10^5 particles per cubic centimeter; however, one would suspect that the actual density is probably less than 100 particles/cm³ in most cases. A density of 50 particles/cm³ would be enough to explain most of the observed magnetic effects of the cloud at the orbit of the earth. We shall now examine the origin of this explosive outburst of gas in more detail from the point of view of hydrodynamics. In the model of the corona which we constructed we could calculate the asymptotic value of the velocity of the solar wind at the orbit of Earth for any particular temperature specified. A few results are compiled in the Table 1. We see that temperatures of the corona greater than 3 million degrees give velocity magnitudes comparable to the velocity of the gas clouds produced by a flare. This suggests that the observational phenomena following a flare can be explained by analyzing the effects of transient disturbances of an expanding corona. Temperatures as high as 4 million degrees have been estimated from coronal line width observations following a flare. Radio observations indicate temperatures perhaps as high as 10 million degrees.

Table 1

Coronal temp. in millions °K	Asymptotic velocity in thousands of km/sec
1	0.5
3	1.0
4	1.2

We shall now discuss the propagative of a blast wave in a spherical atmosphere. It is assumed that some disturbance heats the solar corona

in the vicinity of the active region containing the flare to something of the order of 4×10^6 K. The superheated portion of the corona explodes outward into space. From Table 1, the velocity is over 1000 km/sec, so that presumably it would be greater than the steady-state expansion velocity of the corona of 300 km/sec and the velocity of sound ~ 200 km/sec. Thus these velocities will be ignored in comparison to the shock velocity.

We shall also assume that the quiet-day particle densities in interplanetary space before the explosion drop off with distance r as $1/r^2$.

Before starting the formal calculations, we should like to make the following remark about the physical structure of blast waves from the sun. The actual gas cloud observed at the earth's orbit is not composed of the particles present in the explosion in the corona, but rather of those particles picked up by the blast wave as it moves outward through the solar wind. Therefore the initial cloud of gas arriving at the earth in the blast wave is not the cloud of gas that left the corona at the time of the flare.

From the equations of mass, energy, and momentum conservation, we may write down three equations which express the relationship between the velocity, pressure, and densities before (subscript 1) and after (subscript 2) the discontinuous shock transition. These equations are independent of the nature of the transition and are called the Rankine-Hugoniot relations. These relations are:

$$\rho_1 u_1 = \rho_2 u_2 \quad (13.3)$$

$$\rho_1 u_1^2 + p_1 = \rho_2 u_2^2 + p_2 \quad (13.4)$$

$$\frac{1}{2} \rho_1 u_1^3 + u_1 \left(\frac{f}{2} + 1 \right) p_1 = \frac{1}{2} \rho_2 u_2^3 + u_2 \left(\frac{f}{2} + 1 \right) p_2 \quad (13.5)$$

where

$$1/2f = \frac{\gamma}{\gamma-1} - 1$$

We shall assume that the velocity u_1 is so high that we may neglect the thermal velocity of the incoming gas. We, therefore, ignore the terms

containing p_1 . Under these assumptions one gets the following relations:

$$u_2 = u_1 \frac{\gamma - 1}{\gamma + 1} \quad (13.6)$$

$$p_2 = \frac{2}{\gamma + 1} \rho_1 u_1^2 \quad (13.7)$$

These equations are valid in the frame of reference moving with the shock front. If V is the velocity of the shock front then in a stationary frame in which the gas in front of the shock wave has zero velocity, the gas behind moves with velocity

$$u = u_1 - u_2 = V \frac{2}{\gamma + 1} \quad (13.8)$$

Furthermore,

$$p = \frac{2}{\gamma + 1} \rho_1 V^2 \quad (13.9)$$

$$\rho_2 = \rho_1 \frac{\gamma + 1}{\gamma - 1} \quad (13.10)$$

We see that the density is bounded by $4\rho_1$ regardless of the fact that the velocity is not bounded. Increasing the velocity of the shock merely heats the gas so that its compressibility is limited. In our analysis we are going to compute the asymptotic form of the blast wave at large r , assuming that it will be descriptive of the blast wave, to some degree, at the orbit of Earth. The equations for the blast wave are the following:

Mass conservation,

$$\frac{\partial \rho}{\partial t} + \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 \rho u) = 0 \quad (13.11)$$

Momentum conservation,

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial r} = \frac{1}{\rho} \frac{\partial p}{\partial r} \quad (13.12)$$

Consistent with the assumption of an adiabatic process, we have

$$\frac{\partial p}{\partial t} + u \frac{\partial p}{\partial r} = \left(\gamma \frac{p}{\rho} \right) \left(\frac{\partial \rho}{\partial t} + u \frac{\partial \rho}{\partial r} \right) \quad (13.13)$$

$$\frac{\partial \mathcal{E}}{\partial t} + \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 u \mathcal{E}) = 0 \quad (13.14)$$

where

$$\mathcal{E} = \frac{1}{2} \rho v^2 + \frac{p}{\gamma - 1} = \text{energy density} \quad (13.15)$$

Ignoring the thermal energy ahead of the shock, the total energy of the blast wave is just

$$E = \int_0^R r^2 dr \left(\frac{1}{2} \rho v^2 + \frac{p}{\gamma - 1} \right) \quad (13.16)$$

The following similarity form allows us to reduce this set of partial differential equations to a set of ordinary differential equations:

$$\eta = \frac{t}{r^\lambda} \quad (13.17)$$

where we let

$$\rho = \frac{\Omega(\eta)}{r^2} \quad u = \frac{r}{t} v(\eta) \quad p = \frac{P(\eta)}{t^2} \quad (13.18)$$

The profile of the blast wave is described by η , and we assume that in the asymptotic solution everything is a function of η . The complete details of this treatment may be found in "Supersonic Flow and Shock Waves."²¹

The total energy written as

$$E(t) = - \frac{t(3/\lambda) - 2}{\lambda} \int \frac{d\eta}{\eta(3/\lambda) + 1} \left(\frac{1}{2} \Omega v^2 + \frac{p}{\gamma + 1} \right) \quad (13.19)$$

To illustrate the meaning of the parameter λ , suppose that we have a short-lived heating of the solar corona to 4×10^6 K. The energy of the blast wave is then constant with time and we have $\lambda = 3/2$. This tells us that the velocity of the wave will decrease as it moves outward. In other words as it expands and gathers more particles its velocity must reduce to conserve energy. However, in the sun the high temperature of 4 million degrees observed for a flare sometimes lasts as much as a day, which means that energy may perhaps be supplied to the blast wave by the pressure of the corona for some time. If we take the extreme case of $\lambda = 1$, we have the energy increasing linearly with the time which means the blast wave will then have a constant velocity. In this extreme case we have a scooping up of interplanetary gas toward the rear of the blast wave due to the continuing pressure from the corona. For a 4 million degree corona, a rough estimate gives about 30 to 80 particles/cc at the head of the blast wave when it reaches the orbit of the earth (approximately four times the quiet-day density), whereas the density at the rear of the blast wave may be a few hundred particles/cm³ if the enhanced corona pushes firmly from behind. The expected blast wave velocities at this temperature turn out to be 1 or 2 thousand km/sec. We would expect the actual situation to lie between $\lambda = 1$ and $\lambda = 3/2$ even though the calculations for $\lambda = 3/2$ seem to be more consistent with observation.²²

In the case of $\lambda = 3/2$ we can obtain algebraic solutions of the equations which allow us to solve explicitly for all parameters. In this case we find that the density varies linearly behind the shock wave. The solution is illustrated in Fig. 13.1. The profile jumps up at most a factor of 4 in crossing the shock wave.

In the case when λ is less than $3/2$ we expect that the additional energy fed into the wave will simply scoop up the tail of the blast wave as is illustrated in Fig. 13.1. The point that is being illustrated is that the blast wave in both cases first scoops up the quiet-day interplanetary material and forces it out in its expansion. Since the gas carries the interplanetary magnetic fields with it, the magnetic field immediately

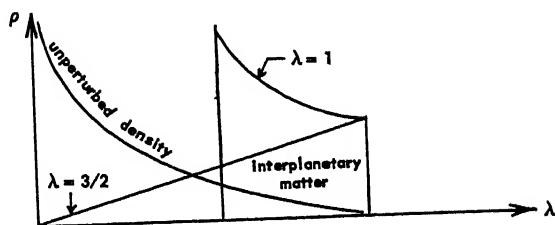


Fig. 13.1

behind the wave will be the compressed interplanetary field, and the magnetic field from the flare site could only appear at Earth after this initial portion of the wave has passed.

We shall now briefly indicate how to go about solving these equations analytically. We put the similarity solutions into the hydrodynamical equations and then define a new variable

$$C^2 = \gamma \frac{P}{\Omega} \quad (13.20)$$

This is essentially a measure of the speed of sound to an observer who moves so that η is constant.

The conservation equations become after considerable reduction

$$\begin{aligned} [(1 - \lambda U)^2 - \lambda^2 C^2] \eta \frac{dU}{d\eta} &= U(1 - U)(1 - \lambda U) \\ + \frac{C^2}{\gamma} [2\lambda - 3\gamma\lambda U] \end{aligned} \quad (13.21)$$

$$\begin{aligned} 2[(1 - \lambda U)^2 - \lambda^2 C^2] \frac{\eta}{C} \frac{dC}{d\eta} &= 2 + U(1 - 3\lambda - 3\gamma + \lambda\gamma) \\ + 2\gamma\lambda U^2 + \frac{2\lambda^2 C^2}{1 - \lambda U} \left(U - \frac{1}{\gamma} \right) \end{aligned} \quad (13.22)$$

$$[(1 - \lambda U)^2 - \lambda^2 C^2] \frac{\eta}{P} \frac{dP}{d\eta} = 2 + U(\gamma\lambda - 3\lambda - 3\gamma) + 2\gamma\lambda U^2 \quad (13.23)$$

The problem has been reduced somewhat in that we now have a set of ordinary differential equations to solve rather than partial differential equations. The procedure from here on is straightforward. We divide the first equation by the second to obtain an equation of the form dU/dC . The Rankine-Hugoniot relations then give us boundary conditions from which we can start integrating at the head of the shock backward, toward the rear. These R. H. relations are

$$P(\eta_1) = \frac{2\rho_c}{\lambda^2(\gamma+1)} \quad (13.24)$$

$$\Omega(\eta_1) = \frac{\gamma+1}{\gamma-1} \rho_c \quad (13.25)$$

$$U(\eta_1) = \frac{2}{\lambda(\gamma+1)} \quad (13.26)$$

And so we obtain U as a function of C behind the shock wave. We note that λ and γ are constants in the problem. We need a value of U at the rear of the blast wave in order to know where to cut the integration off. We obtain this by the following method. The velocity is

$$\frac{dR_2}{dt} = \frac{t^{(1/\lambda)-1}}{\lambda\eta_2^{1/\lambda}} = \frac{t^{1/\lambda}}{\eta_2^{1/\lambda}} U(\eta_2) \quad (13.27)$$

Since we know that this derivative must equal the original definition of the velocity, equating the two at η_2 gives us the proper asymptotic time dependence of U

$$U(\eta_2) = \frac{1}{\lambda} \quad (13.28)$$

We then stop the integration when U reaches this value. We now have U as a function of C and so may solve the second equation and obtain C as a function of η . Similarly the third equation can be solved for P . These solutions when carried out exhibit the forms illustrated above.

Time will not allow us to consider the subjects of cosmic rays, aurorae, or magnetic storms. These subjects are related very closely to the topics we have discussed up to this point.

We have attempted in this series of lectures to trace out the passage of energy and material from the convective zone of the sun into the corona and from there out into space. Our discussion has been limited to a rather qualitative presentation. We have for the most part ignored such effects as cosmic rays, the rotation of the sun, and anisotropies in the corona, in order to obtain a simple over-all picture in the time allowed to us for this presentation of the principal effects.

Many of the assumptions made in the treatment of these topics are still to be verified by observation. Space vehicles should, in the near future, give much more accurate measurements of such things as the densities and magnetic fields in interplanetary space, and, therefore, allow us to evaluate better many of the assumptions made at the present time.

BIBLIOGRAPHY AND NOTES, Secs. 1 to 6

- (1) T. G. Cowling, "Magnetohydrodynamics," Interscience, New York, 1956.
- (2) L. Spitzer, "Physics of Fully Ionized Gases," Interscience, New York, 1955.
- (3) Chandrasekhar and Trehan, "Plasma Physics," Univ. of Chicago Press, Chicago, 1960.

The above books are good general reference books on the formulation and treatment of plasmas.

- (4) K. M. Watson, Phys. Rev., 102, 12 (1956). Discusses static solution and gives individual particle motions in the guiding center approximation.

(5) K. A. Brueckner and K. M. Watson, Phys. Rev., 102, 19 (1956). Discusses dynamic solution for small radius of gyration. Uses moment equations. They express the perturbation f' in the distribution function as:

$$f^1 = A_1 (C_{\perp}^2, C_1^2, r) + C_1 A_2 (C_{\perp}^2, C_1^2, r)$$

where C_1 is the particle velocity along the field and C_{\perp} is the velocity perpendicular to the field. They give two coupled linear equations for A_1 and A_2 with an inhomogeneous term $G [(d\xi/dt), C_{\perp}^2, C_1^2]$ which is the Boltzmann operator of f' . The pressure tensor is diagonal in the local Cartesian system in all these cases of guiding center motion.

(6) Chew, Goldberger, and Low, Proc. Roy. Soc. (London), A236, 112 (1956).

They take the moment equations and obtain

$$\begin{aligned} \rho_0 \frac{d\bar{u}_0}{dt} = & -\nabla P_0 + \frac{(\nabla \times \bar{B}) \times \bar{B}}{4\pi} + \frac{\partial}{\partial t} \frac{(\bar{V}_0 \times \bar{B}) \times \bar{B}}{C^2} \\ & + \frac{(\bar{V}_0 \times \bar{B}) \cdot \nabla \cdot (\bar{V}_0 \times \bar{B})}{C^2} \end{aligned}$$

where the last two terms are small in the usual case that

$$\frac{B^2}{8\pi} \ll \rho C^2$$

They also have

$$\frac{\partial \bar{B}}{\partial t} = \nabla \times (\bar{V}_0 \times \bar{B}), \quad \frac{\partial \rho_0}{\partial t} = -\nabla \cdot (\rho_0 \bar{V}_0)$$

where the subscript zero denotes the unperturbed motion, i.e., the \bar{V}_0 is the electric drift, ρ_0 is the unperturbed density. The drifts of order (R/L) are neglected explicitly in the equation. They deduce that if the

heat transport is neglected along the field, then

$$\frac{d}{dt} \left(\frac{P_{\parallel} B^2}{\rho_0} \right) = 0 \quad \frac{d}{dt} \left(\frac{P_{\perp}}{\rho_0 B} \right) = 0$$

where $d/dt = (\partial/\partial t) + \bar{V}_0 \cdot \nabla$. For slow variation along the field we have $B \propto \rho_0$ which from the first relation tells us that

$$\frac{\partial}{\partial t} \left(\frac{P_{\parallel}}{\rho_0} \right) = 0 \quad \text{or} \quad \frac{\partial \Gamma_{\parallel}}{\partial t} = 0$$

The second relation is the usual adiabatic invariant.

(7) Chandrasekhar, Kaufman, and Watson, Proc. Roy. Soc. (London), A245, 435 (1958). This article discusses stability of the pinch effect and gives a formal derivation of hose instability and mirror instability. Noting page 455, one has

$$\eta = \frac{P_{\perp}}{P_{\parallel}} \quad \beta = \frac{4\pi}{B^2} P_{\perp}$$

For the mirror instability let $V = 0$, i. e., $k_{\parallel} = 0$. Then (A9) becomes

$$2\beta(1 - \eta) + 1 < 0$$

$$\frac{P_{\perp}}{B^2/8\pi} \left(1 - \frac{P_{\perp}}{P_{\parallel}} \right) + 1 < 0$$

$$1 - \frac{P_{\perp}}{P_{\parallel}} + \frac{B^2}{8\pi P_{\perp}} < 0$$

which gives instability if ρ_{\parallel} is sufficiently small. This is mirror instability.

For the hose instability, let $\theta = \pi/2$, i. e., $k_{\perp} = 0$ and

(A9) becomes

$$\beta \left(1 - \frac{1}{\eta}\right) + 1 = 0$$

$$1 - \frac{p_{\parallel}}{p_{\perp}} + \frac{B^2}{4\pi p_{\perp}} = 0$$

$$p_{\perp} - p_{\parallel} + \frac{B^2}{4\pi} < 0$$

which gives instability if p_{\parallel} is sufficiently large. Note that, if $p_{\perp} = B^2/4\pi$, then the mirror instability criterion reduces to $p_{\perp} > 2p_{\parallel}$. For hose instability, $p_{\parallel} < 2p_{\perp}$. If $B^2/4\pi \ll p_{\perp}$, then the criteria are $p_{\parallel} < p_{\perp}$, $p_{\parallel} > p_{\perp}$, respectively.

(8) A. Schluter, Z. Naturforsch., 5a, 72 (1950); Ann. Physik, 10, 422 (1952). Discusses the role of conductivity in the equations of a collisionless plasma.

References

1. E. N. Parker, Phys. Rev., 107, 924-933 (1957).
- 1a. G. P. Kuiper (ed.), "The Sun," Univ. of Chicago Press, Chicago, 1953.
2. G. Bahng and M. Schwarzschild, Astrophys. J., 134, 312 (1961).
3. A. Pellew and R. V. Southwell, Proc. Roy. Soc. (London), A173, 312, (1961).
- 3a. W. V. R. Malkus, Proc. Roy. Soc. (London), A225, 185, 196 (1954).
4. S. Chandrasekhar, "Hydrodynamic and Hydromagnetic Stability," Oxford, London, 1961.
5. P. Ledoux, M. Schwarzschild, and E. A. Spiegel, Astrophys. J., 133, 184 (1961).
6. M. J. Lighthill, Proc. Roy. Soc. (London), 211, 564 (1952); 222, 565 (1954).
7. R. M. Kulsrud, Astrophys. J., 121, 461 (1955).
8. E. N. Parker, Phys. Rev., 90, 240 (1953).
9. P. M. Morse, "Vibration and Sound," McGraw-Hill, New York, 1948.

10. I. Proudman, Proc. Roy. Soc. (London, 214, 119 (1953).
11. D. Osterbrock, Astrophys. J., 134, 347 (1961).
12. W. A. Whitaker, Astrophys. J., in press.
13. S. Chapman, Astrophys. J., 120, 151 (1954).
14. S. R. Pottasch, Astrophys. J., 131, 68 (1960).
15. E. N. Parker, Astrophys. J., 132, 821 (1960).
16. J. W. Chamberlin, Astrophys. J., 133, 675 (1961).
17. F. S. Clauser, 4th Symp. on Sosomal Electrodynamics, IAU, Varenna, Italy, August 1960.
18. L. Biermann, Observatory, 77, 109 (1957).
19. A. J. Deutsch, Astrophys. J., 123, 210 (1956).
20. H. D. Babcock, Astrophys. J., 130, 364 (1959).
21. R. Courant and K. O. Friedrichs, "Supersonic Flow and Shock Waves," Sec. 160, Interscience, New York, 1948.

STELLAR EVOLUTION

Jack S. Goldstein
Brandeis University

CONTENTS

Introduction	85
1. Observational Material and Descriptive Estimates	86
2. Physical Estimates for Main Sequence Stars	87
3. Formation of the Protostar	88
4. Stellar Evolution of the Main Sequence	96
5. Main Sequence Stars and Post-Main Sequence Development	99
6. Degenerate Configurations	105
References	106

INTRODUCTION

In these three lectures, we shall attempt to trace as much of the life history of a star as is presently possible, from its presumed birth to its supposed death, owing to the exhaustion of its nuclear energy supply. It is clear that some of the mathematical arguments will have to be presented in rather sketchy form and that some of the physical arguments will necessarily receive less than their just due; but where this is necessary, I shall attempt to provide sufficient references to the literature so that the interested reader may pursue these matters further.

1. OBSERVATIONAL MATERIAL AND DESCRIPTIVE ESTIMATES

It seems fair to begin with a brief description of the observational evidence which must be fitted by any theory of stellar evolution. Detailed calculations, to which we shall have later reference, show quite clearly that the time scales associated with different stages in the evolution of a star may be different by many orders of magnitude; thus, the period of hydrogen burning, for a star of roughly solar characteristics, is expected to be of the order of 10^{10} to 10^{11} years, but may be as short as 10^6 years for the most massive stars; the period of deuterium burning, through which a protostar may pass before reaching the hydrogen burning stage, is probably of the order of 10^6 years, or less; whereas the final, cataclysmic stages of those stars that end as supernovae are probably to be measured in seconds or fractions thereof.

It is evident that we have a better chance of observing stars while they are in their longer lived stages of development, and it is true that most (but by no means all) of the stars that are observed are in the hydrogen burning period.

The fundamental data that form the basis for theory and conjecture can be described most readily by the Hertzsprung-Russell diagram. Here, on logarithmic scales, the absolute luminosity of a star is plotted against its temperature, or against some quantity representative of temperature. In Fig. 1.1, such data are plotted schematically. The curve ABCD represents the majority of observed stars and is known as the main sequence.

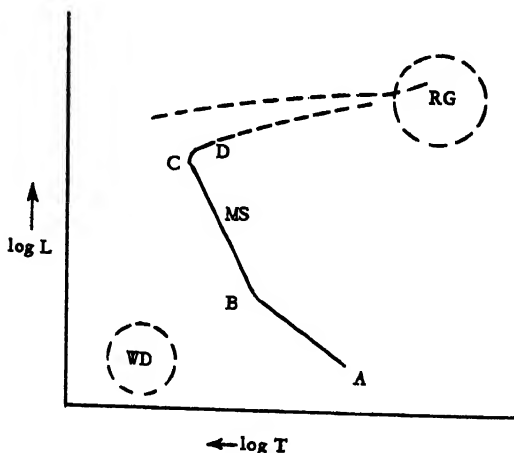


Fig. 1.1

When the H-R diagram is constructed for all stars, independent of association with each other, the main sequence appears as a more or less indistinct band in the diagram rather than as a single sequence. This fact is accounted for by differences in age, chemical composition, etc. On the other hand, if the H-R diagram is constructed for a single association of stars, such as a globular cluster, the main sequence emerges quite clearly. In the case of a globular cluster, we can assume that, since the association was formed at a single epoch, the individual stars have the same age and chemical composition and differ only in mass.

The upper right-hand corner of the H-R diagram represents a class of star known as the red giants, characterized by relatively low surface temperature and high total luminosity (implying a large radius). The lower left corner contains the white dwarfs, characterized by high surface temperature and low luminosity, and, consequently, small radius. The red giants and white dwarfs are believed to represent evolutionary stages of stars that have left the main sequence after burning a critical fraction of their hydrogen supply.

The points B and C, indicated in Fig. 1.1, may be called the turn-on and turn-off points of the main sequence. The significance of these points will be made clear in the subsequent lectures.

It should be made clear that although the H-R diagram does, in fact, contain much of the evolutionary history of a star, the main sequence does not represent an evolutionary track for a star; that is, the motion of a star in the H-R diagram, as it evolves, is not from one end of the main sequence to the other.

2. PHYSICAL ESTIMATES FOR MAIN SEQUENCE STARS

In order for a star to remain on the main sequence for any length of time, it must be in hydrostatic equilibrium. Any departure from hydrostatic equilibrium would induce changes in the star which must occur in times of the order of the free-fall time, $(GM/R^3)^{-1/2}$, which, for the sun, is of the order of only $\sim 10^3$ sec.

The equations of hydrostatic equilibrium are, for a spherically symmetric star,

$$\frac{d}{dr} P(r) = - \frac{GM(r)\rho(r)}{r^2} \quad (1)$$

$$\frac{d}{dr} M(r) = 4\pi r^2 \rho(r)$$

where P and ρ represent local pressure and density and $M(r)$ is the mass interior to the radius r . Simple estimates of the central conditions can be obtained by assuming approximately linear dependence of P and ρ on r ; we immediately find for the central pressure and temperature the approximate values (for the sun)

$$P_c \approx 10^9 \text{ atm} \approx 10^{15} \text{ dynes/cm}^2$$

$$T_c \approx 10^7 \text{ }^\circ\text{K}$$

The estimate of temperature is based on the assumption that the ideal gas law holds and that the star consists mainly of ionized hydrogen.

The rate at which the sun radiates away its energy can be compared with its gravitational and thermal energy stores. It is found that the total gravitational and thermal energy of the sun are sufficient to account for a solar lifetime no longer than about 3×10^7 years; since good geological evidence exists for the fact that the sun has radiated at a quite constant rate for at least one hundred times that long, it is clear that something must happen at 10^7 °K to stabilize the star and to supply the radiated energy.

Evaluating kT at 10^7 °K, we find kT to be of the order of 1 kev; thus, we expect the essential physics to consist of low-energy nuclear reactions in the kilovolt region.

If a mass M of hydrogen is completely converted to helium, an amount of energy $0.007 Mc^2$ is liberated. This amount of energy, radiated away at the observed solar rate, leads to a solar lifetime of the order of 10^{11} years.

3. FORMATION OF THE PROTOSTAR*

If we try to retrace stellar evolution to the earliest possible epoch, we find very few observational facts to guide us. The earliest stages are either nonluminous and hence not observable, or else short lived, and so also not observable. The principal observational material to which we can look for guidance is summarized in the following few statements:

*This material is principally taken from an article by Hoyle (reference 1).

(1) Galaxies appear to occur in clusters, and at least one author considers that there is evidence for "clusters of clusters." Hence, small groups of galaxies are formed in association with each other.

(2) The total mass of a galaxy typically lies between the limits

$$3 \times 10^9 M_{\odot} \leq M_{\text{gal}} \leq 3 \times 10^{11} M_{\odot}$$

(3) The mass of a typical star of population II (those stars which are believed to be of galactic age) is of order M_{\odot}

(4) An appreciable fraction of the mass of a galaxy occurs in the form of interstellar gas and dust. An appreciable fraction of the mass of any large sample of the universe probably occurs in the form of intergalactic gas clouds, presumably of hydrogen.

The smallest density that can be realized by a gas cloud which is to contract under its own gravitational attraction is of the order of 10^{-27} g/cc. This can be seen from the following rough argument. We suppose that any two particles recede from each other according to Hubble's law for the expansion of the universe:

$$v = \frac{1}{T} r \quad (2)$$

where T is Hubble's constant. Differentiating to obtain the rate of change of momentum,

$$\dot{p} = \frac{m}{T} \dot{r} = \frac{m}{T^2} r$$

This rate of change of relative momentum must be balanced by the attractive force due to gravity if contraction is to occur. Assuming, therefore, a spherically symmetric cloud, an elementary volume dV at radius r , experiences a net "force"

$$F = \rho dV \left[\frac{r}{T^2} - \frac{GM(r)}{r^2} \right] \quad (3)$$

where $M(r)$ is the mass internal to radius r ; $M(r) \simeq (4\pi r^3 \rho / 3)$. Hence for contraction

$$\frac{GM(r)}{r^3} \geq \frac{1}{T^2}$$

or

$$\frac{4}{3} \pi \rho G \geq \frac{1}{T^2} \quad (4)$$

Introducing numerical values for G and T , the figure of 10^{-27} g/cc is arrived at for ρ . We note for future use that the contraction time itself, τ , for any gaseous configuration, by which we mean the time to contract, say, to half the size, is of the order of $\tau \sim (\rho G)^{-1/2}$, by simple dimensional arguments.

It is known from radio observations that the intergalactic clouds cannot be more dense than about 10^{-26} g/cm³. Furthermore, mass motions are observed in such clouds with velocities in the range 10 km/sec $< v < 100$ km/sec.

Assuming that this kinetic energy is turbulent, it is eventually converted into heat, and is therefore available as a source of thermal energy, certainly in excess of 10^{12} ergs/gm, and possibly as large as 10^{14} ergs/gm.

Assuming the gas to be atomic hydrogen, the specific heat is such that from 10^{12} to 10^{13} ergs/gm is sufficient to raise the gas to temperatures of the order of 10^4 °K, at which temperature, collisional ionization begins, becoming complete at 2.5×10^4 °K. Thus, at slightly more than 10^{13} ergs/gm, all the hydrogen will be ionized. The gas will remain approximately in this same state for a time of the order of the contraction time, i. e., $\tau \sim (\rho G)^{-1/2} \simeq 10^{17}$ sec.

During this time, a certain amount of energy, E , is radiated away by the gas, which is optically thin at all wavelengths with the possible exception of L_{α} . According to Minkowski,² E is given by

$$E = 1.5 \tau \times 10^{-27} \left(\frac{\rho}{m_H} \right)^2 \sqrt{T} \left[1 + \frac{3.85 \times 10^5}{T} \right] \text{ ergs/cm}^3 \quad (5)$$

The first term represents free-free transitions, the second, free-bound. At 2.5×10^4 °K, the gas is wholly ionized, and L_{α} radiation escapes; Eq. (5) then holds. At 1.0×10^4 °K, L_{α} is imprisoned, and the constant 3.85×10^5 is replaced by 6.5×10^4 .

Thus, at a temperature of 1.5×10^4 °K, we find E to be 3×10^{13} ergs/gm; at 2.5×10^4 °K, E has the value 1.3×10^{14} ergs/gm. These amounts are greater than the available thermal energy and greater than the amounts necessary to ionize all the gas. On the other hand, if no ionization takes place at all, no appreciable amount of energy can be radiated. Consequently, the temperature must equilibrate in the neighborhood of 10^4 °K, unless enough energy is available to ionize all the gas. In that case, the temperature must equilibrate at a temperature near 3.85×10^5 °K. Reference to Fig. 3.1 will make this point clear. The significant feature in this diagram is that the curve is very flat between temperatures of 2.5×10^4 and 3.85×10^5 °K; the chance is small, therefore, that the available energy will have just that value to cause the gas to equilibrate at some temperature between these two values. In any case, on the estimate of the available thermal energy quoted above, it is to be expected that the gas will equilibrate at the lower temperature of about 10^4 °K.

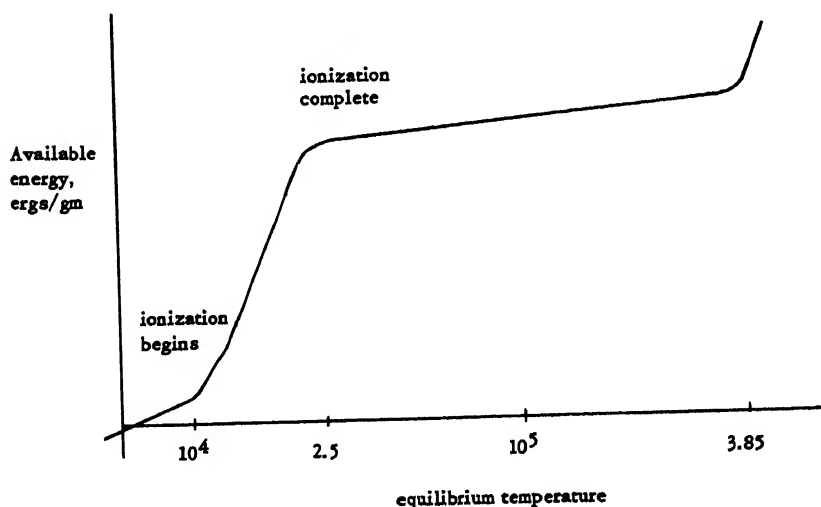


Fig. 3.1

The virial theorem states that in order for a system to contract, the gravitational potential energy must exceed twice the thermal (kinetic) energy. Thus, we require

$$\frac{GM}{V^{1/3}} \gtrsim 5RT \quad (6)$$

where M is the total mass of the gas, V , its volume, T , the temperature, and R , the ideal gas constant. Since the gas is assumed to be only partly ionized, the internal energy per gram is an average between $3/2 RT$ and $3 RT$; the coefficient $5/2$ is taken as the approximate average.

Putting $M = \rho_0 V$, with $\rho_0 = 10^{-27}$, and taking T to be 1.5×10^4 , we find

$$M^{2/3} \geq \frac{5RT}{G\rho_0^{1/3}} = \left(1.4 \times 10^{10} M_\odot\right)^{2/3} \quad (7)$$

We find, therefore, that if a system is to contract at all from an initial density of 10^{-27} g/cm^3 , it must have at least galactic mass. If the temperature is taken to be $1.0 \times 10^4 \text{ }^\circ\text{K}$, the total mass is reduced to $3.6 \times 10^9 M_\odot$.

It is readily seen that as long as the gas is essentially transparent, so that Minkowski's formula [Eq. (5)] applies, the contraction must be isothermal. The time taken to contract to half the original size is of order 10^{17} sec; this, by the arguments already given, is sufficiently long for the cloud to radiate away any excess in energy released by the contraction. A further contraction by half takes place in a time shorter by $\sqrt{2}$; but the rate of radiation increases by a factor of 4 due to the presence of ρ^2 in Minkowski's formula. Hence the argument goes through a fortiori. Until the gas is sufficiently dense so that it becomes partially opaque, contraction must take place essentially isothermally, at a temperature in the neighborhood of $10^4 \text{ }^\circ\text{K}$.

Now the work done in an isothermal contraction of a differential volume, dv , is

$$dW = -p dv = -R\rho T dv = -RT \frac{dv}{v}$$

per gram. Hence, if a mass M is compressed isothermally from volume V_1 to volume V_2 , the total work done is

$$W = MRT \ln \left(\frac{V_1}{V_2} \right) \quad (8)$$

The gravitational energy released is simply

$$\frac{GM^2}{V_1^{1/3}} \left[\left(\frac{V_1}{V_2} \right)^{1/3} - 1 \right] \quad (9)$$

If we assume $M \gg M_{\min}$, where M_{\min} is given by Eq. (7), we can write

$$\frac{GM^2}{V^{1/3}} \gg 5MRT \quad (10)$$

Thus, for any appreciable volume ratio, the gravitational energy release is much larger than the increase in thermal energy.

For the contraction to be quasi-static, however, the virial theorem requires that half of the gravitational energy go into thermal energy; otherwise, the additional energy must go into mass motions which serve to re-expand the cloud.

It follows, therefore, that the gas cloud cannot contract quasi-statically and uniformly unless some other mechanisms exist for the dissipation of gravitational energy. Energy that appears in the form of shock waves or turbulence must go eventually into heat; and we are led to the consideration that subcondensations may occur.

If, in fact, this does happen, we find that the contraction time is unchanged, since it depends only on initial density; but the gravitational potential energy released by each subcondensation is now smaller, since this term depends on $M^2/V^{1/3}$; hence it is possible to have approximate equality between the gravitational and thermal energies in each of the subcondensations. It is evident that for such equality, the mass of the subcondensation must be of order M_{\min} .

Now, the ratio V_1/V_2 cannot much exceed 3 before the inequality re-expresses itself. The dependence of the gravitational energy on this ratio is much stronger than that of the thermal energy.

We assume the following numerical values: $M_{\min} = 3.6 \times 10^9 M_{\odot}$; $\rho_{\odot} = 10^{-27}$; and $V_1/V_2 = 3$. We then find that the contracted size of the subcondensation is approximately 13,000 pc, which is roughly the observed size of the galactic halo.

Once the contraction proceeds to the point where the gravitational and thermal energies are no longer of the same magnitude, a new set of

subcondensations must be formed in order for further contraction to take place.

As a simple mathematical model, suppose that the system contracts by a factor q in volume, and then divides into k equal masses. If R_0 represents the original radius, the k equal masses will be characterized by

$$\text{radius} = \frac{R_0}{(kq)^{1/3}} \quad \text{density} = \rho_0 q \quad \text{mass} = \frac{M_0}{k}$$

If each set of subcondensations bears a constant relation to the previous set, we find that the contraction times form a sequence

$$\tau, \tau q^{-1/2}, \tau q^{-1}, \tau q^{-3/2}, \dots \quad (11)$$

or that the total time occupied by n steps is

$$\tau_{\text{total}}^{(n)} = \tau(1 + q^{-1/2} + q^{-1} + \dots) = \tau \frac{1 - q^{-n/2}}{1 - q^{-1/2}} \quad (12)$$

If we take $q = 3$, the total time necessary to go through, say, ten stages is essentially the time necessary to go through an infinite number of stages; and

$$\tau_{\text{total}}^{(\infty)} = 2.3\tau \quad (13)$$

The final stages, of course, occur very quickly.

We must now ask what can happen to prevent this sequence from winding up with all matter in the form of microstars. The condition for isothermal contraction is necessarily violated as soon as the gas becomes partially opaque. Complete opacity would mean, in essence, that the contraction is adiabatic. Under this circumstance, the thermal energy per gram is (assuming the ratio of specific heats to be $5/3$)

$$-\int p \, dv = -KR \int v^{-5/3} \, dv$$

where we use the adiabatic relation

$$T = K V^{-(\gamma-1)}$$

Hence

$$W = \frac{3}{2} R M T_0 \left[\left(\frac{V_1}{V_2} \right)^{2/3} - 1 \right] \quad (14)$$

The dependence on the volume ratio is now stronger than the dependence of the gravitational energy; consequently, the gas mass, which is now to be called a protostar, can satisfy the virial theorem and continue to contract until the conditions for the release of nuclear energy are met at its center. At that point, the additional pressure either from radiation or from convection is sufficient to stabilize the star.

Now, in terms of our simplified model, the gravitational energy release in the n -th step is

$$\begin{aligned} \frac{G M_n^2}{V_n^{1/3}} \left[\left(\frac{V_n}{V_{n+1}} \right)^{1/3} - 1 \right] &= \frac{G M_0^2 (k q)^{n/3}}{k^{2n} V_0^{1/3}} [q^{1/3} - 1] \\ &= \frac{G M_0^2}{V_0^{1/3}} \frac{q^{n/3}}{k^{5n/3}} (q^{1/3} - 1) \end{aligned} \quad (15)$$

The time for the n -th contraction is just

$$\tau_n = (G \rho_n)^{-1/2} = \tau_0 q^{-n/2}$$

If L represents the luminosity of the protostar, the condition for complete adiabaticity is clearly

$$L \tau_n << \frac{G M_n^2}{V_n^{1/3}} \left[\left(\frac{V_n}{V_{n+1}} \right)^{1/3} - 1 \right]$$

or

$$L < \frac{GM_0^2}{V_0^{1/3}} \frac{q^{n/3}}{k^{5n/3}} (q^{1/3} - 1)(G\rho_0)^{1/2} q^{n/2} \quad (16)$$

or, taking the approximate equality, and grouping all the numerical factors together into one grand constant

$$L \sim \frac{G^{3/2} M_0^2 \rho_0^{1/2}}{V_0^{1/3}} Q \quad (17)$$

where Q is a number not too different from unity. Equation (17) thus gives approximately the luminosity of a protostar in terms of the initial conditions leading to its formation.

4. STELLAR EVOLUTION TO THE MAIN SEQUENCE*

Once a protostar is formed, its evolution toward the main sequence (M. S.) can be calculated. We shall make certain simplifying assumptions for heuristic purposes, but the only essential assumption is that the contraction to the M. S. may be considered quasi-static.

Consider a protostar of low density (e. g., a radius of $\sim 10R_\odot$) and low central temperature ($T_c < 10^5$ °K). If we assume an initial value for T_c , then, in principle, we may calculate the run of temperature and density (ρ), as well as the radius itself. For this calculation, we use the equations of hydrostatic equilibrium [Eq. (1)]; the local equation of state relating P , T , and ρ ; and the equation of energy transports which, for the case of radiative equilibrium, takes the form

*See E. Salpeter, reference 3, for a more detailed treatment as well as extensive references.

$$\frac{d}{dr} (aT^4) = -\frac{3}{c} \kappa \rho \frac{L(r)}{4\pi r^2} \quad (18)$$

where a is Stephan's constant, and κ , the mean opacity, and c , the velocity of light.

Thus, from the calculated model we obtain the luminosity, L , and the total thermal and gravitational energies, E_{th} and E_{gr} . If models such as this are carried out for a sequence of increasing values of T_c , then this sequence of models may be thought of as representing the quasi-statically contracting protostar.

We may introduce a time scale by writing

$$L = -\frac{d}{dt} (E_{th} + E_{gr}) = -\frac{d}{dT_c} (E_{th} + E_{gr}) \frac{dT_c}{dt} \quad (19)$$

Thus, we can express, e.g., the radius or any other property as a function of time. We find, in fact, that a star spends relatively little of its time contracting, compared to the time spent on the M. S. Consequently, few stars are found in the pre-M. S. stage. The lower part of the Herz-sprung-Russell diagram is shown in Fig. 4.1.

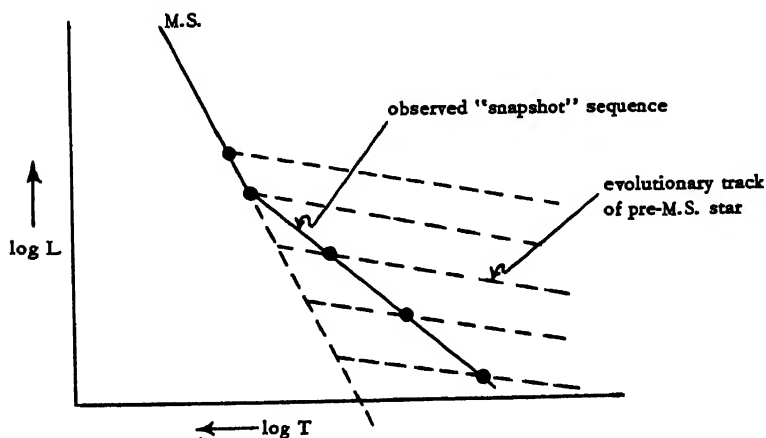


Fig. 4.1

Now, qualitative results may be obtained from the assumption that the contraction is a homologous one, i. e., that the sequence of models differ only by a change in scale. We see at once that the density must scale as

$$\rho \sim \frac{1}{R^3} \quad (20)$$

The equation of hydrostatic equilibrium [Eq. (1)] may now be used to determine that the pressure must scale as

$$P \sim \frac{1}{R^4} \quad (21)$$

and, since

$$P = \frac{k}{\mu} \rho T \quad (22)$$

where μ is the mean molecular weight, the temperature must scale as

$$T \sim \frac{1}{R} \quad (23)$$

One may show that Eq. (19) may be put in the form (after differentiation)

$$\frac{dL}{dr} = -6\pi r^2 \rho^{5/3} \frac{d\left(\frac{P}{\rho^{5/3}}\right)}{dt} \quad (24)$$

Upon introducing the assumption of a homologous transformation, as expressed by Eqs. (20) to (23), this relation becomes

$$\frac{dL}{dr} = -6\pi r^2 P \frac{d}{dt} \log R \quad (25)$$

Thus, the equations of hydrostatic equilibrium, together with Eqs. (18) and (25), lead to a relatively simple model which can be solved by

standard techniques (provided a reasonable assumption is made about the temperature dependence of κ).

The results may be summarized as follows:

(a) The product $L \sqrt{R}$ remains approximately constant throughout the contraction

(b) The contraction time (time to reach the M. S. values of L and R) is given by

$$\tau = \frac{3}{D} \frac{GM^2}{LR} \quad (26)$$

where D is a constant of order unity, but quite sensitive to the opacity law. The contraction time is essentially the gravitational energy divided by the luminosity—a ratio known as the Kelvin time. For a star of 1 solar mass, this time is about 10^7 years.

Since the ratio M/R is roughly constant along the M. S., we obtain the rough law

$$\frac{\tau}{\tau_{\odot}} = \frac{M}{M_{\odot}} \frac{L_{\odot}}{L} \quad (27)$$

For a typical bright (upper M. S.) star of class BO, we find

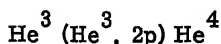
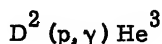
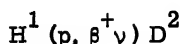
$$\tau \sim 10^5 \text{ years}$$

5. MAIN SEQUENCE STARS AND POST-MAIN SEQUENCE DEVELOPMENT

We shall not discuss the structure of main sequence models, but shall instead discuss the mechanisms for the release of nuclear energy. The reader interested in problems of structure is referred to The Structure and Evolution of the Stars by Schwarzschild.⁴

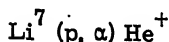
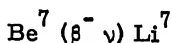
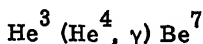
A star consisting only of hydrogen can release nuclear energy in essentially one way, namely, by the proton-proton cycle. This process,

which sets in at a temperature of $\sim 10^7$ °K is the following:



The total energy release is 26.2 Mev/cycle, or 6.3×10^{18} ergs/gm, with 0.52 Mev/cycle carried away by the two neutrinos.

According to Cameron, * when sufficient He^4 has accumulated, this cycle may be modified as follows:



This modification may be more efficient than the original cycle.

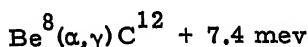
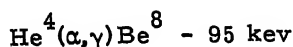
If C^{12} is present in the stellar core, hydrogen burning may proceed by the so-called carbon-nitrogen cycle, which will be dealt with below.

When a sufficient amount of He^4 has accumulated, it acts as a damper for the p-p cycle, and insufficient energy can be generated to support the outer portions of the star. The star must, therefore, contract until the central temperature rises sufficiently for He burning to begin. At this point the star leaves the main sequence, and the stellar configuration consists of a very hot core and a greatly expanded envelope. The star is now in the upper right (red giant) region of the H-R diagram.

He burning is extremely sensitive to temperature, and sets in abruptly at some value of temperature between 1.0 and 1.5×10^8 °K. The

*See acknowledgment at the end of this section.

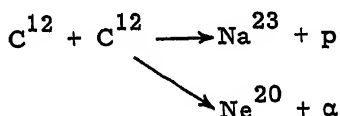
process, known as the 3α reaction, is simply



Although the first step is endothermic, the net energy balance of this reaction is positive, as indicated.

Since C^{12} is easily destroyed by (α, γ) reaction to yield O^{16} , Hoyle postulated that in order to account for the observed (terrestrial + solar + meteorite) abundance of C^{12} , the 3α process must be characterized by a resonance. He predicted the spin and parity of the resonance and an energy of 7.68 Mev. The resonance was subsequently found at 7.65 Mev.

Thus, the end result of helium burning is C^{12} and O^{16} . At a temperature of about 0.8 BK, * we can have



where the proton/alpha is immediately consumed.

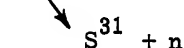
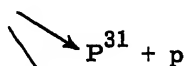
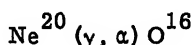
By (α, γ) reactions, then, the group of elements from Ne to Mg is formed. Each process is characterized by a threshold value of temperature, which is reached by a contraction of the star whenever the previous process reaches exhaustion. According to the calculations of Cameron, and of Chiu, at a temperature of 0.5 to 0.6 BK, a significant population of electrons and positrons is produced by the tail of the Planck distribution. Occasionally (if the weak interaction theory is correct), positron-electron annihilation results in a $(\nu - \bar{\nu})$ pair:



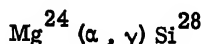
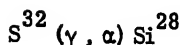
*1 BK = 10^9 °K.

The importance of this process lies in the extreme transparency of the star to neutrinos. Once a neutrino pair is formed, it escapes at once, and the star must compensate for this energy loss by further contraction. But, by the virial theorem, half the energy so released goes into raising the temperature, and, since the reaction rates are extremely sensitive to temperature,* all processes are speeded up enormously.

Now, at 1.5 to 2 BK, (γ , α) reactions can take place, and the first of these is



and at ~ 2 BK,



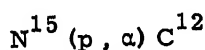
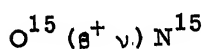
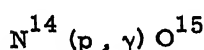
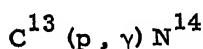
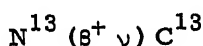
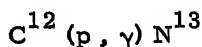
There is thus a tendency at this stage for the abundances to cluster around Si^{28} .

At 3 to 4 BK, all Coulomb barriers can be overcome, and we reach the Fe-Ni peak, with the abundances of all the elements determined by considerations of statistical equilibrium. Once this stage is reached, the further evolution of the star must be very rapid. The gravitational contraction no longer results in a further production of energy, but, in fact, results in the "ionization" of iron nuclei into helium nuclei. Since this requires about 1.6×10^{18} ergs/gm, and the entire gravitational energy store

*The 3α reaction rate goes approximately as T^{30} , in the neighborhood of 0.2 BK.

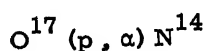
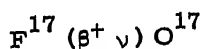
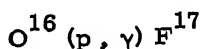
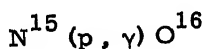
is only of the order of 0.3 Mev/gm, the star continues to contract increasingly quickly, and with an essentially isothermal core because of electron degeneracy. The neutrino loss is now very severe, and the contraction extremely rapid. Presumably the inward falling, outer layers, still containing unburned nuclear fuel, are rapidly heated, and the star explodes in a supernova outburst. Needless to say, these final stages of evolution are highly conjectural.

If now new stars are formed from interstellar gas which has been mixed with the products of nucleogenesis, the new stars will have some C^{12} present in the initial stages of hydrogen burning. If the star is massive enough to require a high luminosity, the p-p cycle may be replaced by the C-N cycle:

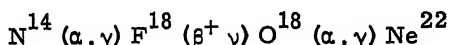


Since the p-p reaction rate is dominated by essentially temperature independent nuclear factors, it is constrained to be slow, whereas the C-N reaction can go much faster if the temperature is raised slightly. Consequently, stars more massive than about 1 solar mass burn hydrogen mostly by the C-N process. Both the C-N and p-p processes go at about equal rates in the sun.³

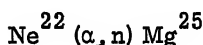
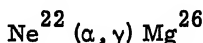
Now, the C-N cycle has a side loop that is possible. The last step may be replaced by



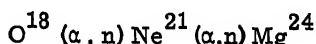
The N^{14} is returned to the C-N cycle, in consequence of which the abundance of N^{14} will be anomalously high. This makes possible the chain



which is multiply resonant. Ne^{22} , at appropriate temperatures, may react with α particles:



The second reaction is slightly endothermic, but may, nevertheless, represent an important source of neutrons which are needed for the buildup of heavy elements by capture by the Fe-Ni group. There are, of course, many reactions that may take place during the supernova explosion itself if sufficient helium is present; thus, for example:



For further details on the buildup of heavy elements, the reader is referred to the remarkable article by Burbidge et al.⁵

The role of neutrino emission in stellar evolution is discussed by H. Y. Chiu, in several publications of The Goddard Institute for Space Studies, New York.*

Many of the reaction mechanisms discussed above have been suggested by A. Cameron, and are reported with his kind permission from a lecture given by him in 1962. However, Dr. Cameron has not had opportunity to evaluate this manuscript, and any errors must be considered to be due to the present writer.

*See also Chiu, reference 6.

6. DEGENERATE CONFIGURATIONS

When the density in the stellar core becomes sufficiently high, electron degeneracy sets in, and the pressure of the electron gas must be considered. It was shown by Chandrasekhar⁷ that a configuration with mass less than $1.44 M_{\odot}$ can be stable even without energy generation. Consequently, the evolutionary scheme leading to the supernova, described above, requires stars of mass in excess of the Chandrasekhar limit.

It can be shown fairly directly that a critical mass should exist, although its actual value can only be obtained from a detailed calculation. The pressure of a degenerate electron gas, in the nonrelativistic limit, is given by

$$P \propto \rho^{5/3} \quad (28)$$

Whereas in the relativistic limit,

$$P \propto \rho^{4/3} \quad (29)$$

where we neglect the contribution from finite temperature. Thus, arguing from a dimensional analysis of the equation of hydrostatic stability, we have

$$\frac{P_c}{R} \sim K \frac{GM^2}{R^5} \quad (30)$$

where P_c represents the central pressure, and K represents an approximate constant. Introducing Eq. (28), for the nonrelativistic case, we have, approximately

$$\frac{M^{5/3}}{R^6} \sim K' \frac{GM^2}{R^5} \quad (31)$$

or

$$M^{1/3} R \simeq \text{constant.} \quad (32)$$

Thus, we find that the radius of a nonrelativistic degenerate configuration decreases with increasing mass; however, as long as the configuration remains nonrelativistic, there always exists a nonvanishing radius for every mass.

If the mass is sufficiently large, however, the density, and consequently the Fermi energy, must increase to the point where relativistic degeneracy is reached. In that case, Eq. (31) is altered to

$$M^{4/3}/R^5 \sim K'GM^2/R^5 \quad (33)$$

It is seen that the radius drops out of the equation of hydrostatic balance, and for a sufficiently large mass, the right-hand side (representing the gravitational contraction) must exceed the left-hand side (representing the pressure gradient). The configuration must therefore collapse for some critical value of the mass.

A full treatment of very dense configurations may be found in the articles of references 8, 9, and 10.

References

1. F. Hoyle, *Astrophys. J.*, 118, 513 (1953).
2. R. Minkowski, *Astrophys. J.*, 96, 206 (1942).
3. E. Salpeter, *Revs. Modern Phys.*, 29, 244 (1957).
4. M. Schwarzschild, The Structure and Evolution of the Stars, Princeton Univ. Press, Princeton, New Jersey, 1958.
5. Burbidge, Burbidge, Fowler, and Hoyle, *Revs. Modern Phys.*, 29, 547 (1957).
6. H. Y. Chiu, *Ann. Phys.*, 15, 1 (1961); *ibid.*, 16, 321 (1961).
7. S. Chandrasekhar, Introduction to Stellar Structure, Univ. of Chicago Press, Chicago, Illinois, 1939.
8. E. Salpeter, *Energy and Pressure of a Zero-Temperature Plasma*, *Astrophys. J.*, 134, 669 (1961).
9. T. Hamada and E. Salpeter, *Models for Zero-Temperature Stars*, *Astrophys. J.*, 134, 683 (1961).
10. V. Ambartsumyan and G. Saakyan, *The Degenerate Superdense Gas of Elementary Particles*, *Soviet Astronomy-AJ (English Transl.)*, 4, 187 (1960).

Note Added in Proof: The theory of the fragmentation of gas clouds, due to Hoyle, has been called into serious question in a pair of articles by D. Layzer, to appear in the Astrophysical Journal in 1963. An alternate mechanism is proposed by Layzer, based only on the fact that the universe is expanding. A modified virial theorem, which is concerned only with the energy of density fluctuations, indicates that such fluctuations should grow and become more bound as time increases.

TOPICS IN THE THEORY OF THE VIBRATIONS
OF IMPERFECT CRYSTALS

A. A. Maradudin
Westinghouse Research Laboratories
Pittsburgh, Pennsylvania

CONTENTS

I. Introduction	110
II. A One-Dimensional Example	114
1. Vibrations of a Perfect Linear Chain	114
2. A Mass defect in a Linear Chain	119
III. General Theory for Isolated Defects	129
IV. A Mass Defect in a Three-Dimensional Crystal	153
1. Properties of the Normal Modes	153
2. Thermodynamic Functions of a Cubic Bravais Crystal Containing a Single Mass Defect	166
3. Example: The Second-Order Doppler Shift in the Energy of a Recoilless Gamma-Ray Photon	173
V. Thermodynamic Functions of Isotopically Disordered Crystals	185
1. Introduction	185
2. The Helmholtz Free Energy of an Isotopically Disordered Crystal	189
VI. Frequency Spectrum of an Isotopically Disordered Crystal	208
1. Introduction	208
2. The Frequency Spectrum in the Limit of Low Impurity Concentration	209
3. Machine Calculations of the Frequency Spectrum of a Disordered Linear Chain	221
VII. The Phonon Propagator	226
1. The Phonon Propagator for a Disordered Crystal	226
2. The Phonon Propagator for a Bravais Crystal Containing a Single Mass Defect	248

CONTENTS (Continued)

VIII. Optical Absorption in Isotopically Disordered Crystals	253
1. Formula for the Optical Absorption Coefficient	253
2. The Absorption Coefficient and the Propagator Formalism	258
3. Results for the Absorption Coefficient	261
IX. The Mössbauer Effect	272
1. The Absorption Cross Section	272
2. The Impurity Debye-Waller Factor	280
3. The One-Phonon Absorption Cross Section	285
X. Lifetime of a Localized Vibration Mode	293
1. Introduction	293
2. The Lifetime of a Localized Vibration Mode	299
3. Numerical Estimate of the Lifetime of a Localized Mode	308
References	317

I. INTRODUCTION

It is perhaps expressing a truism to say that few solids in nature are free from defects and imperfections of some type. This fact has, for some time, been recognized and taken into account in the development of several branches of solid-state physics—notably those dealing with the mechanical or plastic properties of solids and in semiconductor physics.

In this respect, lattice dynamics has lagged behind other branches of solid-state physics. The study of the effects of defects and disorder on the vibrational properties of solids is of comparatively recent origin. The reason for this may lie in the fact that it took lattice dynamics some 20 years to escape from the Babylonian captivity of the Debye approximation, and, in the first flush of freedom, its practitioners proceeded first to calculate vibration spectra for a wide variety of solids and then to consider in detail the thermal and optical properties on the basis of progressively more realistic models. This kind of work, which is still going on

today, has been significantly aided in recent years by the development of neutron spectroscopy and the discovery of the Mössbauer effect, which provide methods for the experimental study in great detail of the dynamics of crystal lattices.

To my knowledge, the earliest studies of the effect of defects and disorder on the vibrational properties of crystals were those carried out by I. M. Lifshitz and his collaborators in Russia beginning in 1943 and continuing up to about 1956.¹ Their earliest papers, published during the war in relatively inaccessible journals, in a rather inaccessible language, went largely unnoticed by other workers in the field until virtually all their work had been redone, simplified, and extended. A similar program of research was undertaken by Montroll and his collaborators at the University of Maryland in 1955,² and more recently Litzman in Czechoslovakia³ and Hori⁴ and Teramoto⁵ in Japan have been actively engaged in this kind of work.

Studies of the effects of defects and disorder on the vibrational properties of crystals can be divided into two not-always-distinct categories: those which deal with the effects of defects and disorder on equilibrium properties of a crystal, and those which deal with their effects on dynamical properties. Into the former category fall the calculations of the discrete frequencies associated with the spatially localized vibration modes from isolated impurity atoms and the calculations of the changes in additive functions of the normal mode frequencies, such as the thermodynamic functions, resulting both from isolated impurities and from finite concentrations of impurities. Roughly speaking, this category is composed of those problems that require explicit or, more usually, implicit knowledge of the frequency spectrum of the perturbed crystal.

Into the second category fall those calculations that deal with the effects of defects and disorder on the motion of individual atoms or in which individual normal modes play the dominant role. Examples of problems of this type are the calculations of the optical absorption in mixed or isotopically disordered alkali-halide crystals and of the second-order Doppler shift in the energy of a γ -ray emitted in a recoilless transition by a nucleus bound in a crystal.

Of these two kinds of problems, I think it is fair to say that the first is by far the more extensively studied and the one for which a rather well-developed theory exists. It is now possible to work out any of the thermodynamic functions of a three-dimensional two-component isotopically randomly disordered crystal for a wide range of mass ratios and compositions without explicit knowledge of the frequency spectrum of the disordered crystal over as much of the temperature range, as it is feasible to do such calculations for a perfect crystal.

It is also possible to evaluate, by numerical methods, the frequencies and atomic displacements in the localized vibration modes associated with isotopic impurities for realistic models of three-dimensional crystals.

The extension of such calculations to take account of force-constant changes as well as mass changes is hampered at present only by computational complications. With the advent of high-speed computers it has become possible to calculate by purely numerical means the frequency spectra of disordered two-component linear chains with nearest-neighbor interactions.⁶ Dean and his co-workers at the National Physical Laboratory in England have already discussed the extension of these machine methods to the calculation of the spectra of disordered two-dimensional lattices,⁷ and I am sure that it will not be long now before such calculations are carried out for three-dimensional lattices.

I do not want to give the impression that all the problems of the equilibrium theory of imperfect lattices have been solved. The effects of extended imperfections, such as dislocations, on the vibrations of crystals have not yet been treated quantitatively to any great extent, and there is a whole host of unsolved problems in the equilibrium theory of disordered crystals. Some of the more interesting of these are those in which correlations in the positions of the defects are considered. A recent calculation by Kirkwood and Wojtowicz⁸ of the vibrational contribution to the specific heat anomaly in β -brass at the critical temperature is the first significant contribution to the study of this kind of problem. Interesting problems also arise when the simultaneous effects of defects and lattice anharmonicity are considered. However, the more interesting of these, for example the calculation of the lifetime of a localized vibration mode,⁹ and the determination of the lattice thermal conductivity of isotopic mixtures,¹⁰ fall more properly into the category of dynamical problems.

Thus, in summary, the main problems in the equilibrium theory of the vibrational properties of imperfect lattices are well understood, some of them have already been tackled successfully, and the main areas in the subsequent development of this subject will probably consist of calculations based on more realistic models of crystals, and on more realistic models of impurities and defects, and of generalizations such as have just been mentioned.

In contrast to this reasonably favorable stage of development in which the equilibrium theory finds itself, the dynamical theory of imperfect crystals, which is of more recent origin, is rather less well developed. It is a theory, however, that deserves to be developed more fully at the present time. The recent emergency of x-ray diffuse scattering techniques, neutron spectroscopy, and the Mössbauer effect, as sensitive probes of the dynamics of individual atoms, combined with theoretical analyses of these dynamical properties should yield results of interest not only to the theory of lattice dynamics, but, as I shall try to show in these lectures, to other branches of physics as well.

Indeed, although one of my aims in giving this series of lectures is

to present the elements of an exact theory of the effects of defects and disorder on the vibrational properties of crystals, more importantly, I should like to apply the results of the theory to the explanation of several recent experiments in which localized vibration modes due to impurities play an important role, and to the discussion of possible new experiments that would show up the effects of impurities on the vibrations of crystals.

The outline of these lectures is as follows: After a brief discussion of some defect problems for a linear chain, in which case many problems can be solved exactly, we present the elements of the theory of lattice dynamics for crystals containing few or isolated defects. This case corresponds to problems that are of interest in themselves, and also provides a set of results that are useful in calculations of the effects of a large number of defects on the vibrational properties of crystals. We then turn to a study of the vibrations of randomly disordered crystals and describe ways of calculating equilibrium properties of such crystals, such as the thermodynamic functions and the frequency distribution function. We also show how one can calculate time-dependent properties of disordered crystals as well and illustrate them by obtaining expressions for the optical absorption coefficient of a disordered crystal, and the cross section for the absorption of γ -rays by an impurity atom in a crystal. We conclude by presenting a calculation of the lifetime of a localized impurity mode due to its coupling with the other modes of the crystal through anharmonic forces.

The present survey is not intended to be exhaustive. In particular, we discuss only briefly the interactions between defects through their vibration fields and the scattering of sound waves by defects. The former problem is discussed, for example, in the papers quoted in Ref. 2. The latter problem has been discussed by Lifshitz,¹¹ and a more recent study by Koster¹² of the scattering of conduction electrons by defects in a crystal is also relevant to this problem. In addition, we do not discuss explicitly the time dependence of the individual normal coordinates of a disordered crystal. This problem is discussed in Ref. 13, and some of the discussion in Sec. VII of these notes also bears on it. Finally, we remark that since even the calculation of the simplest vibrational property of a perfect crystal, its frequency spectrum, requires extensive numerical work, it is not surprising that the more difficult calculations that I shall be describing should lead to more difficult numerical computations. Accordingly, in the applications of the theory to particular examples we shall quite frequently adopt cubic Bravais crystals as the unperturbed crystals and shall sometimes assume a Debye frequency spectrum. These two assumptions will enable us to carry out analytically calculations which otherwise would be intractable without the aid of high-speed computers.

II. A ONE-DIMENSIONAL EXAMPLE

1. Vibrations of a Perfect Linear Chain

In this section we introduce some of the techniques we shall employ in the remainder of these lectures by applying them to a simple example. We study here some of the effects on the vibrations of a linear chain that are produced when one of its atoms is replaced by an impurity whose mass differs from that of the atom it replaces. Even in one dimension this problem has all the features which are found in its three-dimensional analog, but it can be discussed without the complications of notation and computation that necessarily attend the latter problem. It also gives us results to which we can refer at corresponding points in the subsequent more general analysis.

We begin by reviewing briefly some properties of the vibrations of a perfect linear chain.

The Hamiltonian for a linear chain of N atoms interacting with nearest neighbor forces is

$$H = \frac{M}{2} \sum_{n=-\frac{N}{2}+1}^{\frac{N}{2}} \dot{x}_n^2 + \frac{1}{2} \gamma \sum_{n=-\frac{N}{2}+1}^{\frac{N}{2}} (x_n - x_{n-1})^2 \quad (2.1.1)$$

where $x_n(t)$ is the displacement of the n -th atom from its equilibrium position at time t , M is the atomic mass, and γ is the nearest neighbor atomic force constant — the second derivative of the interatomic potential function evaluated at the nearest neighbor separation.

In writing Eq. (2.1.1), for reasons of mathematical convenience, we have adopted the cyclic boundary condition on the atomic displacements. This condition can be expressed as

$$x_n = x_{n+N} \quad (2.1.2)$$

In the limit of very large N the adoption of this rather artificial boundary condition leads to a negligible error in calculations of bulk properties of the crystal and also in calculations of properties associated with individual atoms.

Since $p_n = M \dot{x}_n$ is the momentum conjugate to the displacement x_n , the equations of motion of the atoms of the crystal can be written

$$\dot{p}_n = - \frac{\partial H}{\partial x_n}$$

or

$$M \ddot{x}_n = \gamma(x_{n+1} - x_n) - \gamma(x_n - x_{n-1})$$

$$n = -\frac{N}{2} + 1, \dots, \frac{N}{2} \quad (2.1.3)$$

The time independent equations of motion of the linear chain are obtained from Eq. (2.1.3) by making the substitution $x_n = u_n e^{i\omega t}$, where u_n is independent of time. The equations satisfied by the $\{u_n\}$ are

$$M\omega^2 u_n + \gamma [u_{n+1} - 2u_n + u_{n-1}] = 0$$

$$n = -\frac{N}{2} + 1, \dots, \frac{N}{2} \quad (2.1.4)$$

The solutions of this set of equations can be written

$$u_n = e^{in\phi} \quad (2.1.5)$$

provided that the frequency ω is related to the phase angle ϕ by

$$M\omega^2 = 2\gamma - 2\gamma \cos \phi \quad (2.1.6)$$

The values of ϕ are not arbitrary. According to the cyclic boundary the following condition must be satisfied:

$$e^{in\phi} = e^{i(n+N)\phi}$$

so that

$$\phi = \frac{2\pi s}{N} \quad s = -\frac{N}{2} + 1, \dots, \frac{N}{2} \quad (2.1.7)$$

where s is an integer. The restriction of s to the interval $(-N/2) + 1, N/2$ is due to the fact that, if we replace s by s plus any multiple of N , we merely reproduce the solution that corresponds to s in this interval. All the distinct solutions to our problem are thus found in the interval $(-N/2) + 1 \leq s \leq N/2$, and we see there are as many solutions as there are degrees of freedom in the chain. The ϕ -interval $(-\pi, \pi)$ in which all the distinct solutions lie is called the first Brillouin zone for our simple crystal.

From Eqs. (2.1.6) and (2.1.7) we see that the allowed frequencies of the linear chain are given by

$$\omega_s = \left| \omega_L \sin \frac{\pi s}{N} \right| \quad s = -\frac{N}{2} + 1, \dots, \frac{N}{2} \quad (2.1.8)$$

where $\omega_L = (4\gamma/M)^{1/2}$ is the largest frequency of our one-dimensional crystal. The solutions of Eq. (2.1.4) corresponding to the frequency ω_s are doubly degenerate since they are given by the two linearly independent expressions $u_n(s) = \exp(2\pi i s n/N)$ and $u_n(-s) = \exp(-2\pi i s n/N)$.

By taking suitable linear combinations of these solutions, we can construct solutions having either even or odd parity about a given atom.

The frequencies $\{\omega_s\}$ are called normal mode frequencies of the chain. The name derives from the fact that we can perform a principal axis transformation on the Hamiltonian of the chain to reduce it to a sum of N linear harmonic oscillator Hamiltonians, each with its own characteristic frequency. The set of the characteristic frequencies is the set of the $\{\omega_s\}$. Since we shall make frequent use of this transformation in what follows, we outline it briefly for a one-dimensional crystal.

The principal axis transformation is based on the fact that the functions $\{N^{1/2} \exp(2\pi i s n/N)\}$ form a complete orthonormal set. They satisfy the direct and inverse orthonormality conditions

$$\sum_n \left\{ \frac{1}{\sqrt{N}} e^{\frac{2\pi i s n}{N}} \right\} \left\{ \frac{1}{\sqrt{N}} e^{\frac{2\pi i s' n}{N}} \right\}^* = \Delta(s - s') \quad (2.1.9a)$$

$$\sum_s \left\{ \frac{1}{\sqrt{N}} e^{\frac{2\pi i s n}{N}} \right\} \left\{ \frac{1}{\sqrt{N}} e^{\frac{2\pi i s' n}{N}} \right\}^* = \delta_{nn'} \quad (2.1.9b)$$

where $\Delta(s)$ equals unity if $s = 0, \pm N, \pm 2N, \dots$, and vanishes otherwise. In the present case, since both s and s' are restricted to the interval $(-N/2) + 1, N/2$, we can replace $\Delta(s - s')$ in Eq. (2.1.9a) by the usual Kronecker symbol $\delta_{ss'}$.

We now expand x_n and \dot{x}_n as

$$x_n = \frac{1}{\sqrt{NM}} \sum_s e^{\frac{2\pi i s n}{N}} Q_s \quad (2.1.10a)$$

$$\dot{x}_n = \frac{1}{\sqrt{NM}} \sum_s e^{\frac{2\pi i s n}{N}} \dot{Q}_s \quad (2.1.10b)$$

Because x_n and \dot{x}_n are real, Q_s and \dot{Q}_s must satisfy the conditions

$$Q_{-s} = Q_s^* \quad \dot{Q}_{-s} = \dot{Q}_s^* \quad (2.1.11)$$

If we substitute these expansions into Eq. (2.1.1) and use Eq. (2.1.9a), we obtain for the transformed Hamiltonian

$$H = \frac{1}{2} \sum_s \left\{ \dot{Q}_s^* \dot{Q}_s + \omega_s^2 Q_s^* Q_s \right\} \quad (2.1.12)$$

where ω_s is given by Eq. (2.1.8)
From the Lagrangian,

$$L = \frac{1}{2} \sum_s \left\{ \dot{Q}_s^* \dot{Q}_s - \omega_s^2 Q_s^* Q_s \right\}$$

we find that the momentum conjugate to Q_s^* is

$$P_s = \frac{\partial L}{\partial \dot{Q}_s^*} = \dot{Q}_s \quad (2.1.13)$$

so that P_s^* and Q_s are conjugate variables. Hamilton's equations,

$$\begin{aligned} \dot{Q}_s &= \frac{\partial H}{\partial P_s^*} \\ -\dot{P}_s &= \frac{\partial H}{\partial Q_s^*} = \omega_s^2 Q_s \end{aligned}$$

give us the equation of motion of the coordinate Q_s :

$$\ddot{Q}_s + \omega_s^2 Q_s = 0 \quad (2.1.14)$$

In the theory of small vibrations the coordinates $\{Q_s\}$ are called normal

coordinates, essentially because each of them is independent of all the others. There are as many normal coordinates as there are degrees of freedom in the crystal. Each normal coordinate describes an independent mode of vibration with only one frequency, and such vibration modes are called normal modes. When the crystal is vibrating in one of its normal modes every atom in the crystal is vibrating with the same frequency and a phase which varies from atom to atom according to the factor $\exp(2\pi i \mathbf{n} \cdot \mathbf{s}/N)$. In general, we see from Eq. (2.1.10a) that the motion of the crystal as a whole is a superposition of normal mode motions.

After these preliminaries let us see what happens when we disturb the perfect periodicity of our crystal by introducing a defect into it.

2. A Mass Defect in a Linear Chain

Let us denote by L the difference operator operating on u_n to produce Eq. (2.1.4). Now suppose that because of the presence of defects in the lattice some force constants and masses are different from the others. Then the coefficients of certain of the u 's are different from those given by Eq. (2.1.4). The new equation which shows the effects of these local disturbances is

$$L u_n = \sum_k C_{nk} u_k \quad (2.2.1)$$

where the constants C_{nk} characterize the defects. For example, when we change the mass of the particle at the origin of our lattice ($n = 0$) from M to M' and leave the force constants unchanged, C_{nk} vanishes unless $n = k = 0$, and the coefficient C_{00} is given by $(M - M')\omega^2$. In all that follows we always describe the mass of an impurity atom M' in terms of the mass of the atom it replaces by the parameter ϵ which is defined by $\epsilon = 1 - (M'/M)$. The coefficients C_{nk} for the case of a mass defect can be expressed compactly as

$$C_{nk} = \epsilon M \omega^2 \delta_{n0} \delta_{k0} \quad (2.2.2)$$

The most convenient way of solving Eq. (2.2.1) is through the use of

Green's functions. We introduce the function g_{mn} which is the solution of

$$Lg_{mn} = \delta_{mn} \quad (2.2.3)$$

where δ_{mn} is the Kronecker symbol. We immediately verify that the solution to Eq. (2.2.1) is given by

$$u_n = \sum_{\ell, k} g_{n\ell} C_{\ell k} u_k \quad (2.2.4)$$

since an application of the operator L to both sides of this equation with the aid of Eq. (2.2.3) leads to Eq. (2.2.1).

An expression for g_{mn} can readily be obtained as follows. By writing δ_{mn} as

$$\delta_{mn} = \frac{1}{N} \sum_{s=-\frac{N}{2}+1}^{\frac{N}{2}} e^{\frac{2\pi i s}{N}(m-n)} \quad (2.2.5)$$

and g_{mn} as

$$g_{mn} = \frac{1}{N} \sum_{s=-\frac{N}{2}+1}^{\frac{N}{2}} f_s e^{\frac{2\pi i s}{N}(m-n)} \quad (2.2.6)$$

we obtain the following expression for f_s on substituting Eqs. (2.2.5) and (2.2.6) into Eq. (2.2.3):

$$f_s = \left[Mw^2 - 2\gamma + 2\gamma \cos \frac{2\pi s}{N} \right]^{-1}$$

We thus have

$$g_{mn} = \frac{1}{N} \sum_{s=-\frac{N}{2}+1}^{\frac{N}{2}} \frac{e^{\frac{2\pi i s}{N} (m-n)}}{Mw^2 - 2\gamma + 2\gamma \cos \frac{2\pi s}{N}} \quad (2.2.7)$$

It is easily shown (by replacing s by $-s$ as a summation variable) that g_{mn} depends on m and n only through the magnitude of their difference. Because of this it is often convenient to write it as a function of only one index, g_l , where $l = |m - n|$.

It is also important for subsequent applications of the theory being described here to note that the matrix formed from the elements g_{mn} is the inverse of the matrix of the coefficients of the $\{u_n\}$ in Eq. (2.1.4). This conclusion follows directly from the defining equation for g_{mn} , Eq. (2.2.3). The matrix of the coefficients of the $\{u_n\}$ in Eq. (2.1.4) will be denoted by $M_0(\omega)$ in what follows.

The sum appearing in Eq. (2.2.7) can be evaluated in the following way. If we put

$$1 - \frac{Mw^2}{2\gamma} = \cos \phi \quad 0 < \phi < \pi \quad (2.2.8)$$

the expression for g_0 becomes

$$g_0 = \frac{1}{2\gamma N} \sum_{s=-\frac{N}{2}+1}^{\frac{N}{2}} \frac{1}{\cos \frac{2\pi s}{N} - \cos \phi} \quad (2.2.9)$$

Logarithmic differentiation with respect to ϕ of the relation¹⁴

$$\cos N\phi - 1 = 2^{N-1} \prod_{s=-\frac{N}{2}+1}^{\frac{N}{2}} \left[\cos \phi - \cos \frac{2\pi s}{N} \right] \quad (2.2.10)$$

leads to the result

$$g_0 = \frac{1}{2\gamma} \frac{\cot \frac{N\phi}{2}}{\sin \phi} \quad (2.2.11)$$

The general solution of the difference equation (2.2.3) for $n \neq 0$ can be expressed

$$g_n = A \cos n\phi + B \sin |n|\phi \quad (2.2.12)$$

If we set $n = 0$ in this result, we see that $A = g_0$, whereas from Eqs. (2.2.7) and (2.2.12) we find

$$g_1 = g_{-1} = g_0 \cos \phi + \frac{1}{2\gamma}$$

so that $B = (2\gamma \sin \phi)^{-1}$. Combining these results we obtain finally that

$$g_n = \frac{1}{2\gamma \sin \phi} \left[\cot N \frac{\phi}{2} \cos n\phi + \sin |n|\phi \right] \quad (2.2.13)$$

We now apply this method to a simple example, namely the change of the mass of a single atom without any change in the force constants. We suppose that the impurity atom is located at the origin so that the coeffi-

cients C_{nk} are given by Eq. (2.2.2). It is clear that only the symmetric modes, i. e., those for which $u_n = u_{-n}$, are affected by the presence of this defect, since the antisymmetric modes ($u_n = -u_{-n}$) have a node at the position of the impurity and hence do not feel the mass change.

Equation (2.2.4) reduces to

$$u_n = \epsilon M \omega^2 u_0 g_n \quad (2.2.14)$$

Since this equation holds for all n it must hold for $n = 0$. The new normal mode frequencies are obtained by setting $n = 0$ and solving the equation

$$1 = \epsilon M \omega^2 g_0 \quad (2.2.15)$$

for ω^2 . This equation can be rewritten with the aid of Eqs. (2.2.8) and (2.2.11)

$$\tan \frac{N\phi}{2} = \epsilon \tan \phi \quad (2.2.16)$$

The solutions of this equation substituted into Eq. (2.2.8) give the perturbed normal mode frequencies of the lattice. To a first approximation we have

$$\phi_s \approx \frac{2\pi s}{N} + \frac{2}{N} \tan^{-1} \left(\epsilon \tan \frac{\pi s}{N} \right) \quad (2.2.17)$$

$$s = 1, 2, \dots, \frac{N}{2}$$

so that the s -th normal mode frequency is given approximately by

$$\omega_s \approx \omega_L \sin \frac{\pi s}{N} + \frac{\omega_L}{N} \cos \frac{\pi s}{N} \tan^{-1} (\epsilon \tan \frac{\pi s}{N})$$

$$s = 1, 2, \dots, \frac{N}{2}$$
(2.2.18)

Because we are considering only the symmetric modes, which are the only modes affected by the perturbation, we have restricted s to lie in the interval $(1, N/2)$, since we get no new frequencies outside this interval. The number of symmetric modes given by this restriction taken together with an equal number of (unperturbed) antisymmetric modes gives the correct total number of independent normal modes for the lattice.

We see from Eq. (2.2.18) that if the impurity has a mass which is lighter than that of a normal atom ($\epsilon > 0$), the frequencies are increased above their unperturbed values, whereas they are decreased if the impurity mass is heavier than a normal mass.

In the case that $\epsilon > 0$, a solution to Eq. (2.2.16) exists which corresponds to a frequency which is greater than the maximum frequency of the unperturbed lattice. This solution is obtained by replacing ϕ by

$$\phi = \pi + iz$$
(2.2.19)

Equation (2.2.16) becomes (in the large N limit)

$$\epsilon \coth \frac{z}{2} = 1$$
(2.2.20)

and has the solution

$$z = \ln \frac{1 + \epsilon}{1 - \epsilon}$$
(2.2.21)

The frequency of this mode is found from Eqs. (2.2.8), (2.2.19), and (2.2.21) to be

$$\omega^2 = \frac{\omega_L^2}{1 - \epsilon^2}$$
(2.2.22)

Since the time-independent amplitude of vibration of the n -th atom

when the crystal is vibrating in its s -th normal mode is given by

$$u_n(s) \propto e^{in\phi_s} \quad (2.2.23)$$

we see that the displacements are wavelike in the cases described by the in-band solutions, Eq. (2.2.17), but that they die off exponentially with increasing distance from the impurity in the case of the special mode whose frequency lies above the band of allowed frequencies:

$$u_n \propto (-1)^n \left(\frac{1 - \epsilon}{1 + \epsilon} \right)^{|n|} \quad (2.2.24)$$

For this reason such a mode is referred to as a "localized mode."

The preceding results are a special case of a general theorem due to Lord Rayleigh¹⁵ which can be stated: If in a dynamical system composed of an array of masses coupled to each other by Hookeian springs a single mass is reduced by δM , all frequencies are unchanged or increased but by no more than the distance to the next unperturbed frequency. An increase of a single mass by δM leaves the frequencies unchanged or reduces them by amounts no greater than the distance to the next unperturbed frequency. The increase (decrease) of a single force constant has the same effect as the decrease (increase) of a single mass.

Since the shift in the maximum frequency of the lattice in the presence of a light mass impurity is independent of the dimensions of the lattice, whereas the shifts of the in-band frequencies are of $O(N^{-1})$, it follows that in calculations of the effects of defects on additive functions of the normal mode frequencies, the contribution from the localized mode frequency is of the same order as the total contribution from all the in-band frequency shifts.

In addition to the bound state problem we have just discussed we can also study scattering problems by essentially the same formalism. However, we now require a different form for the Green's function, one which corresponds to an outgoing wave. We can obtain the desired form in the following way. We formally pass to the limit $N \rightarrow \infty$ and replace the discrete variable $2\pi s/N$ by a continuous variable and replace summation by integration:

$$g_n = \frac{1}{2\gamma} \frac{1}{2\pi} \int_0^{2\pi} \frac{e^{in\phi}}{\cos \phi - \cos \phi_0} d\phi \quad (2.2.25)$$

where we have put

$$w^2 = w_L^2 \sin^2 \frac{\phi}{2} \quad (2.2.26)$$

The result (2.2.25) is purely formal since we have not as yet given a prescription for dealing with the singularity in the integrand. We now do so with the aid of the boundary condition on the Green's function. We make a change of variable

$$z = e^{i\phi} \quad z_0 = e^{i\phi_0}$$

and obtain

$$g_n = \frac{1}{\gamma} \frac{1}{2\pi i} \oint \frac{z^n}{(z - z_0)(z - z_0^{-1})} dz \quad (2.2.27)$$

where the integral is carried out around the unit circle in a counterclockwise sense.

Let us first consider the case $n \geq 0$. The integrand in Eq. (2.2.27) possesses simple poles at $z = z_0$ and $z = z_0^{-1}$. Since $|z_0| = 1$, these poles lie on the contour of integration, and the value of the integral depends on the way in which we integrate around these singularities. A moment's reflection shows us that in order that the Green's function describe an outgoing wave traveling in the positive x -direction, we must choose our contour so that it includes the pole at $z = z_0$ but excludes the pole at $z = z_0^{-1}$ (see Fig. II.1). According to the residue theorem we obtain for g_n the result

$$g_n = \frac{1}{\gamma} \frac{z_0^n}{z_0 - z_0^{-1}} = \frac{e^{in\phi_0}}{2i\gamma \sin \phi_0} \quad n \geq 0$$

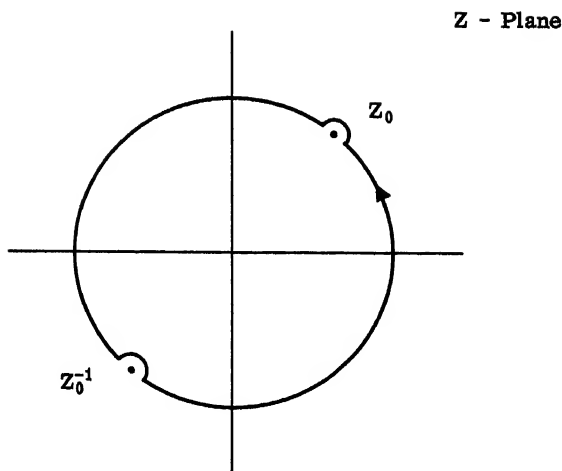


Fig. II. 1 The integration contour for the evaluation of the Green's function integral for scattering problems [Eq. (2. 2. 27)].

In the case that $n \leq 0$ the integral (2. 2. 27) takes the form

$$g_n = \frac{1}{\gamma} \frac{1}{2\pi i} \oint \frac{1}{z^{|n|} (z - z_0)(z - z_0^{-1})} dz \quad (2. 2. 29)$$

and we see that there is an n -th order pole at $z = 0$ in addition to the poles at $z = z_0, z_0^{-1}$. It is readily found that in order that the outgoing wave condition be satisfied we must again choose the contour shown in Fig. II. 1. The result for g_n in this case is found to be

$$g_n = \frac{e^{i|n|\phi_0}}{2i\gamma \sin \phi_0} \quad n \leq 0 \quad (2. 2. 30)$$

We can now consider the problem of the scattering of a wave by a mass defect. We assume that we have an incident wave whose frequency

is ω and which is described by a propagation constant ϕ . In order that this wave be a possible traveling wave in the crystal, according to Eq. (2.1.6) ω and ϕ cannot be independent but must be related by

$\omega^2 = \omega_L^2 \sin^2 \phi/2$. We assume as the solution to Eq. (2.2.1) the sum of the incident wave and an outgoing wave:

$$u_n = e^{in\phi} + c_0 \epsilon M \omega^2 g_n \quad (2.2.31)$$

The coefficient c_0 is determined by substituting Eq. (2.2.31) into Eq. (2.2.1). In this way we obtain

$$\begin{aligned} e^{in\phi} [M\omega^2 - 2\gamma + 2\gamma \cos \phi] + c_0 \epsilon M \omega^2 \delta_{n0} \\ = \epsilon M \omega^2 \delta_{n0} [1 + c_0 \epsilon M \omega^2 g_0] \end{aligned}$$

The first term on the left side of this equation vanishes in view of Eq. (2.1.6), and we find

$$c_0 = \frac{1}{1 + i\epsilon \tan \frac{\phi}{2}} \quad (2.2.32)$$

We can finally express the atomic displacements as

$$u_n = \frac{e^{in\phi}}{1 + i\epsilon \tan \frac{\phi}{2}} \quad n \geq 0 \quad (2.2.33a)$$

$$= e^{in\phi} - \frac{i\epsilon \tan \frac{\phi}{2}}{1 + i\epsilon \tan \frac{\phi}{2}} e^{i|n|\phi} \quad n \leq 0 \quad (2.2.33b)$$

The transmission and reflection coefficients are given by

$$T = \left[1 + \epsilon^2 \tan^2 \frac{\phi}{2} \right]^{-1} = \left[1 + \frac{\epsilon^2 \omega^2}{\omega_L^2 - \omega^2} \right]^{-1} \quad (2.2.34a)$$

$$R = \epsilon^2 \tan^2 \frac{\phi}{2} \left[1 + \epsilon^2 \tan^2 \frac{\phi}{2} \right]^{-1} = \epsilon^2 \omega^2 \times \left[\omega_L^2 - (1 - \epsilon^2) \omega^2 \right]^{-1} \quad (2.2.34b)$$

These coefficients can also be computed in an approximation which is equivalent to the first Born approximation in scattering theory. The results of this calculation are

$$T = 1 - \epsilon^2 \left(\frac{\omega}{\omega_L} \right)^2 \quad (2.2.35a)$$

$$R = \epsilon^2 \left(\frac{\omega}{\omega_L} \right)^2 \quad (2.2.35b)$$

We see that the results obtained by the Born approximation can differ appreciably from the exact results even for moderate values of ω/ω_L .

III. GENERAL THEORY FOR ISOLATED DEFECTS

In the preceding chapter we have described briefly a calculation of the in-band frequencies and the localized mode frequency for a linear chain containing a mass defect. Green's functions methods were used in these calculations and they yielded convenient expressions for the displacement

amplitudes as well. However, calculations of the normal mode frequencies and vibration amplitudes of a crystal containing defects represent only two aspects of equilibrium defect problems in lattice dynamics. There is yet another (of course closely related) problem which has to be discussed. This is the problem of calculating the changes in additive functions of normal mode frequencies as a result of the introduction of defects into crystals. We shall show in this section how the solution of this problem leads also to the normal mode frequencies and amplitudes of the perturbed crystal.

In our discussion of a mass defect in a linear chain we used Green's function methods to obtain the normal mode frequencies and amplitudes of the perturbed crystal. The use of these methods is not essential for the one-dimensional problem. At least two other nonperturbative methods are available for use in this case. However, these methods depend in an essential way on our ability to order points on a line and cannot readily be extended to three-dimensional problems. The use of Green's function techniques, however, is not restricted to one-dimensional problems, and lattices of all dimensions can be treated on an equivalent basis. The use of these techniques in addition provides us with formally exact solutions to the defect problems being studied. And although the price we may have to pay in certain instances for this formal exactness may be the expenditure of heavy numerical computational effort, nevertheless, even in those cases where exact numerical results cannot presently be obtained, the formally exact expressions that we have can serve as the starting point for sensible approximate calculations. Accordingly, we base the treatment of defect problems which follows on the use of Green's function techniques. Since we shall be concerned in this section with formulating a theory valid for arbitrary crystals, we begin by summarizing briefly those elements of the theory of lattice dynamics which will be required in the remainder of these lectures.

We can consider a crystal to be made up of N unit cells each of which contains r atoms. The position vector of the ℓ -th unit cell relative to an origin located at one of the atoms in the crystal is given by

$$\mathbf{x}(\ell) = \ell_1 \mathbf{a}_1 + \ell_2 \mathbf{a}_2 + \ell_3 \mathbf{a}_3 \quad (3.1.1)$$

where $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ are three noncoplanar vectors which are called the primitive translation vectors of the crystal. ℓ_1, ℓ_2, ℓ_3 are three integers which can be positive, negative, or zero, and to which we refer collectively as ℓ . The position vector of the κ -th atom in a unit cell is denoted by $\mathbf{x}(\kappa)$ ($\kappa = 1, 2, \dots, r$), and it is convenient to adopt the convention that $\mathbf{x}(\kappa = 1) = 0$. Thus the position vector of the κ -th atom in the ℓ -th unit cell is given by

$$\mathbf{x}(\ell) = \mathbf{x}(\ell) + \mathbf{x}(\kappa) \quad (3.1.2)$$

Crystals which contain only one atom per unit cell are called Bravais crystals; all others are called nonprimitive crystals.

The atoms of the crystal execute vibrations about the equilibrium positions (3.1.2). If we denote the α -Cartesian component of the displacement of the atom $(\ell\kappa)$ from its equilibrium position by $u_\alpha(\ell\kappa)$, then the Hamiltonian for the vibrating crystal in the harmonic approximation can be written (omitting the energy of the static lattice)

$$H = \frac{1}{2} \sum_{\ell\kappa\alpha} M_\kappa \dot{u}_\alpha^2(\ell\kappa) + \frac{1}{2} \sum_{\substack{\ell\kappa\alpha \\ \ell'\kappa'\beta}} \Phi_{\alpha\beta}(\ell\kappa; \ell'\kappa') \times u_\alpha(\ell\kappa) u_\beta(\ell'\kappa') \quad (3.1.3)$$

M_κ is the mass of the κ -th atom in a unit cell, and the $\{\Phi_{\alpha\beta}(\ell\kappa; \ell'\kappa')\}$ are the general second-order atomic force constants for the crystal. They are the second derivatives of the potential energy with respect to the displacements evaluated in the equilibrium [i. e., all $u_\alpha(\ell\kappa) = 0$] configuration,

$$\Phi_{\alpha\beta}(\ell\kappa; \ell'\kappa') = \left(\frac{\partial^2 \Phi}{\partial u_\alpha(\ell\kappa) \partial u_\beta(\ell'\kappa')} \right)_0 \quad (3.1.4)$$

They are the generalization to three dimensions and an arbitrary range for the interatomic forces of the force constant γ introduced in connection with the one-dimensional model in the last section.

We assume the cyclic boundary condition on the displacements, which in the present case takes the form

$$u_{\alpha}(\ell) = u_{\alpha}(\ell + L) \quad (3.1.5)$$

where L denotes a vector with components (L, L, L) in terms of which $L^3 = N$. In this case, because of resulting invariance of the crystal against a rigid body translation through one of the vectors $x(\ell)$,

$\Phi_{\alpha\beta}(\ell\kappa; \ell'\kappa')$ depends on ℓ and ℓ' only through their difference.

The equations of motion of the crystal are

$$M_{\kappa} \ddot{u}_{\alpha}(\ell) = - \sum_{\ell'\kappa'\beta} \Phi_{\alpha\beta}(\ell\ell'; \kappa\kappa') u_{\beta}(\ell') \quad (3.1.6)$$

If we make the substitution

$$u_{\alpha}(\ell) = v_{\alpha}(\ell) e^{-i\omega t} \quad (3.1.7)$$

the time independent amplitudes $\{v_{\alpha}(\ell\kappa)\}$ satisfy the equation

$$M_{\kappa} \omega^2 v_{\alpha}(\ell) - \sum_{\ell'\kappa'\beta} \Phi_{\alpha\beta}(\ell\ell'; \kappa\kappa') v_{\beta}(\ell') = 0 \quad (3.1.8)$$

which is the generalization of Eq. (2.1.4).

The condition that the set of equations (3.1.8) have a nontrivial solution is that the determinant of the coefficients vanishes

$$| M_{\kappa} \omega^2 \delta_{\ell\ell'} \delta_{\kappa\kappa'} \delta_{\alpha\beta} - \Phi_{\alpha\beta}(\ell\ell'; \kappa\kappa') | = 0 \quad (3.1.9)$$

and the roots of this equation are the normal mode frequencies of the crystal.

If we make the additional substitution

$$v_{\alpha}(\ell) = \frac{v_{\alpha}(\kappa)}{\sqrt{M_{\kappa}}} e^{2\pi i \mathbf{k} \cdot \mathbf{x}(\ell)} \quad (3.1.10)$$

then the eigenvalue equation for the normal mode frequencies can be written

$$\omega^2 v_\alpha(\kappa) = \sum_{\beta\kappa'} D_{\alpha\beta}(\kappa|\kappa') v_\beta(\kappa') \quad (3.1.11)$$

where

$$D_{\alpha\beta}(\kappa|\kappa') = \frac{1}{\sqrt{M_\kappa M_{\kappa'}}} \sum_l \Phi_{\alpha\beta}(\kappa|\kappa', l) e^{-2\pi i \mathbf{k} \cdot \mathbf{x}(l)} \quad (3.1.12)$$

For each value of \mathbf{k} there are $3r$ solutions for ω^2 , and we denote this fact by writing the frequencies as $\omega_j^2(\mathbf{k})$, ($j = 1, 2, \dots, 3r$). To each $\omega_j(\mathbf{k})$ there corresponds an eigenvector $v(\kappa)$, and we indicate this fact by rewriting $v_\alpha(\kappa)$ as $e_\alpha(\kappa | \mathbf{k}, j)$. The eigenvalue equation thus becomes

$$\sum_{\beta\kappa'} D_{\alpha\beta}(\kappa|\kappa') e_\beta(\kappa' | \mathbf{k}, j) = \omega_j^2(\mathbf{k}) e_\alpha(\kappa | \mathbf{k}, j) \quad (3.1.13)$$

The $3r \times 3r$ matrix $D(\mathbf{k})$ whose elements are obtained from $D_{\alpha\beta}(\mathbf{k}; \kappa|\kappa')$ by pairing the indices $(\kappa|\alpha)$ and $(\kappa'|\beta)$ is readily shown to be Hermitian, and the eigenvector components can be constructed so that they satisfy the orthonormality and closure relations

$$\sum_{\alpha\kappa} e_\alpha^*(\kappa | \mathbf{k}, j) e_\alpha(\kappa | \mathbf{k}, j') = \delta_{jj'} \quad (3.1.14a)$$

$$\sum_j e_\alpha^*(\kappa | \mathbf{k}, j) e_\beta(\kappa' | \mathbf{k}, j) = \delta_{\alpha\beta} \delta_{\kappa\kappa'} \quad (3.1.14b)$$

We can also choose them to satisfy the condition

$$e_{\alpha}(\kappa | \begin{smallmatrix} -\mathbf{k} \\ j \end{smallmatrix}) = e_{\alpha}^{*}(\kappa | \begin{smallmatrix} \mathbf{k} \\ j \end{smallmatrix}) \quad (3.1.15)$$

The normal mode frequencies can be shown quite generally to satisfy the condition

$$\omega_j(-\mathbf{k}) = \omega_j(\mathbf{k})$$

For Bravais crystals we suppress the basis index κ in the preceding results. The eigenvectors $\{e(\mathbf{k}j)\}$ are purely real in this case.

Just as in the one-dimensional case the values of the wave vector are not arbitrary but are restricted by the cyclic boundary condition. It is easy to show that the allowed values of \mathbf{k} are given by

$$\mathbf{k} = \frac{1}{L} (h_1 \mathbf{b}_1 + h_2 \mathbf{b}_2 + h_3 \mathbf{b}_3) \quad (3.1.16)$$

$$h_1, h_2, h_3 = 1, \dots, L$$

where $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ are the primitive translation vectors of the reciprocal lattice and are defined by

$$\mathbf{a}_i \cdot \mathbf{b}_j = \delta_{ij} \quad (3.1.17)$$

Equation (3.1.16) shows us that the allowed values of \mathbf{k} are uniformly and densely distributed throughout a unit cell of the reciprocal lattice. It is usually more convenient to restrict them to a more symmetric volume of the reciprocal lattice. This volume, which is called the first Brillouin zone of the crystal, is obtained by drawing vectors from the origin of the reciprocal lattice to all the other lattice points and constructing the planes which are the perpendicular bisectors of these vectors. The smallest volume enclosing the origin and bounded by these planes is the Brillouin zone. It can be shown to be completely equivalent for our purposes to the unit cell of the reciprocal lattice. Note that the Brillouin zone defined here is smaller in volume by a factor of $(2\pi)^{-3}$ than the zone defined in Sec. II-1.

The functions $\{(1/\sqrt{N}) e^{2\pi i \mathbf{k} \cdot \mathbf{x}(\ell)}\}$ satisfy the direct and inverse orthonormality conditions,

$$\sum_{\ell} \left\{ \frac{1}{\sqrt{N}} e^{2\pi i \mathbf{k} \cdot \mathbf{x}(\ell)} \right\} \times \left\{ \frac{1}{\sqrt{N}} e^{2\pi i \mathbf{k}' \cdot \mathbf{x}(\ell)} \right\}^* = \Delta(\mathbf{k} - \mathbf{k}')$$

(3.1.18a)

$$\sum_{\mathbf{k}} \left\{ \frac{1}{\sqrt{N}} e^{2\pi i \mathbf{k} \cdot \mathbf{x}(\ell)} \right\} \times \left\{ \frac{1}{\sqrt{N}} e^{2\pi i \mathbf{k} \cdot \mathbf{x}(\ell')} \right\}^* = \delta_{\ell \ell'}$$

(3.1.18b)

where $\Delta(\mathbf{k})$ equals unity if $\mathbf{k} = 0$ or any translation vector of the reciprocal lattice, and vanishes otherwise. Since \mathbf{k} and \mathbf{k}' are restricted to lie in the first Brillouin zone, the only translation vector of the reciprocal lattice $\mathbf{k} - \mathbf{k}'$ can equal is zero. We could therefore replace the Δ -function in Eq. (3.1.18a) by the Kronecker symbol $\delta_{\mathbf{k}\mathbf{k}'}$.

Since there are N allowed values of \mathbf{k} , and to each of these correspond $3r$ solutions of Eq. (3.1.13), we see that there are as many normal mode frequencies as there are degrees of freedom in the crystal. Since N is large, \mathbf{k} can be regarded as a continuous variable, and summations over \mathbf{k} can be replaced by integrations throughout the volume of the Brillouin zone according to

$$\sum_{\mathbf{k}} \rightarrow V \int_{Bz} d^3k$$

(3.1.19)

where $V = Na_1 \cdot a_2 \times a_3$ is the volume of the crystal.

We now turn to the problem of evaluating an arbitrary additive function of the normal mode frequencies of a perturbed crystal. For convenience let us denote the normal mode frequencies of a perturbed crystal by $\{\omega_j\}$ where the index j runs over all the normal modes. We are thus concerned with the evaluation of functions which have the general form

$$S = \sum_j f(\omega_j) \quad (3.1.20)$$

For example, S will be the Helmholtz free energy if $f(\omega) = kT \ln \{2 \sinh \hbar\omega/2kT\}$. To evaluate S we use a theorem from the theory of contour integration. If the $\{\omega_j\}$ are the roots of an equation which we denote by

$$|M(\omega)| = 0 \quad (3.1.21)$$

then S is given by

$$S = \frac{1}{2\pi i} \int_C f(z) d \ln |M(z)| \quad (3.1.22)$$

where C is any closed, counterclockwise contour which encloses all the zeroes of $|M(z)|$ but none of the poles of $f(z)$.

Since the value of S for a perfect crystal can be calculated by straightforward methods, e. g., through the use of the frequency spectrum, we shall regard it as known, and shall concern ourselves here only with a calculation of the change in S resulting from the presence of defects. To do this we note that the secular determinant in Eq. (3.1.21), whose zeroes are the $\{\omega_j\}$, can be rewritten

$$\begin{aligned}
 |M(w)| &= |M_0(w) + \delta M(w)| = |M_0(w)| |I + M_0^{-1}(w) \\
 &\times \delta M(w)| = |M_0(w)| |\Delta(w)|
 \end{aligned}
 \tag{3.1.23}$$

where $|M_0(w)|$ is the secular determinant of the perfect crystal, and $\delta M(w)$ is the matrix whose elements describe the changes in $M_0(w)$ as a result of the introduction of the defects. For example, in the one-dimensional problem discussed in section 5, the matrix δM is the negative of the matrix $C = (C_{nk})$. In obtaining Eq. (3.1.23) we have used

the fact that the determinant of the product of two matrices is equal to the product of the determinants of the individual matrices. By using the additive property of logarithms, we see from Eqs. (3.1.22) and (3.1.23) that the change in S resulting from defects is given by

$$\Delta S = \frac{1}{2\pi i} \int_C f(z) d \ln |\Delta(z)|
 \tag{3.1.24}$$

The change in an additive function S because of a single defect at some point in a crystal denoted by α can be expressed as

$$\Delta S_\alpha = \frac{1}{2\pi i} \int_C f(z) d \ln |\Delta_\alpha(z)|
 \tag{3.1.25}$$

The quantity ΔS_α is often referred to as the "self- S " of the defect. In an infinite crystal or in a crystal satisfying the cyclic boundary conditions, ΔS_α is independent of α . We can also introduce the "interaction- S " between a pair of defects located at points α and β of the crystal. This is defined as the difference between the value of S for the crystal containing the two defects at α and β and its value when the defects are infinitely separated. The latter quantity is just the sum of the self- S 's for each of the defects. We thus find

$$\Delta S_{\alpha\beta} = \frac{1}{2\pi i} \int_c f(z) d \ln \frac{|\Delta_{\alpha\beta}(z)|}{|\Delta_{\alpha}(z)||\Delta_{\beta}(z)|} \quad (3.1.26)$$

in an obvious notation.

It should be pointed out that the matrix $M(0)$ is not the dynamical matrix for the perturbed crystal, that is, it is not the matrix whose eigenvalues are the normal mode frequencies of the perturbed crystal. The matrix $M(\omega)$ can be written formally

$$M(\omega) = M' \omega^2 - \mathfrak{F}' \quad (3.1.27)$$

where M' is a diagonal matrix whose $(l\kappa\alpha)$ $(l\kappa\alpha)$ element is the mass of the κ -th atom in the l -th unit cell. The matrix \mathfrak{F}' has for its $(l\kappa\alpha)$ $(l\kappa'\beta)$ element the coefficient $\mathfrak{F}'_{\alpha\beta}(l\kappa; l\kappa')$, which contains the effects of any force constant changes produced by the defect. If we divide the elements of every row and column of $M(\omega)$ by the square root of the mass appearing in the corresponding diagonal element and change the sign of each term, the resulting matrix has the form $D - \omega^2 I$, and D is called the dynamical matrix. We accordingly have the following relation:

$$\frac{|M(\omega)|}{|M_0(\omega)|} = \frac{A}{A_0} \frac{|D - \omega^2 I|}{|D_0 - \omega^2 I|} = |I + M_0^{-1}(\omega) \delta M(\omega)| \quad (3.1.28)$$

where A is the product of the masses of the atoms in the perturbed crystal and A_0 is the corresponding quantity for the unperturbed crystal. D_0 is the dynamical matrix for the unperturbed crystal. Since the perturbation matrix δM ordinarily has a simpler form than the perturbation matrix for the dynamical matrix would have, we have developed our

formalism in terms of the matrix $M(\omega)$ rather than in terms of D .

Since the value of the determinant of a matrix is unchanged if we carry out a similarity transformation on the matrix, if we carry out the transformations which diagonalize D and D_0 , respectively, we can rewrite Eq. (3.1.28) as

$$|\Delta(\omega)| = \frac{A}{A_0} \prod_j \frac{(\omega_j^2 - \omega^2)}{(\omega_{0j}^2 - \omega^2)} \quad (3.1.29)$$

where $\{\omega_{0j}\}$ and $\{\omega_j\}$ are the normal mode frequencies of the unperturbed and perturbed crystals, respectively. From Eq. (3.1.29) we obtain the result

$$\begin{aligned} \frac{d}{d\omega} \ln |\Delta(\omega)| &= 2\omega \sum_j \left\{ \frac{1}{\omega^2 - \omega_j^2} - \frac{1}{\omega^2 - \omega_{0j}^2} \right\} \\ &\sim \frac{2}{\omega} \sum_j (\omega_j^2 - \omega_{0j}^2) \quad \text{as } |\omega| \rightarrow \infty \end{aligned} \quad (3.1.30)$$

The result expressed by Eq. (3.1.30) has the following consequence. If, for large $|z|$, $f(z)$ increases no more rapidly than $|z|^\alpha$, where $\alpha < 2$, we can use for C the D-shaped contour consisting of a semicircle of radius R in the right-hand half-plane together with the portion of the imaginary axis in the interval $(-iR, iR)$ (see Fig. III.1). In the limit as $R \rightarrow \infty$, the contribution from the semicircle vanishes in view of Eq. (3.1.30), and we are left with only an integration down the imaginary axis. All of the thermodynamic functions lead to functions $f(z)$ which satisfy the above criterion. When the contour of Fig. III.1 is used, only the even part of $(i\omega)$ contributes to the integral. This result follows from the fact that $|\Delta(\omega)|$ is a function of ω^2 and thus an even function of ω , so that $d/d\omega \ln |\Delta(\omega)|$ is an odd function of ω .

From Eq. (3.1.29) we see that the roots of the equation

$$|\Delta(\omega)| = 0 \quad (3.1.31)$$

are the frequencies of the normal modes which have been altered owing to the introduction of the defects. The normal mode frequencies that are not changed by the presence of the defects are not given by Eq. (3.1.31) because the factors of $\omega_{0j}^2 - \omega^2$ which would give rise to these roots are

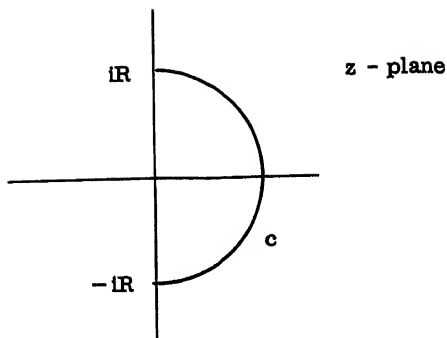


Fig. III. 1 The integration contour in the complex frequency plane which is used in evaluating changes in additive functions of normal mode frequencies.

canceled between the numerator and denominator of the expression on the right side of Eq. (3.1.29). This result provides the first connection between the problems of determining the normal mode frequencies of a perturbed crystal and of calculating changes in additive functions of these frequencies.

A final useful property of the determinant $|\Delta(\omega)|$ is its relation to the change in the frequency spectrum of the crystal resulting from defects. The frequency spectrum $g(\omega)$ is defined in such a way that $g(\omega) d\omega$ is the fraction of normal modes with frequencies in the interval $(\omega, \omega + d\omega)$ in the limit as $d\omega \rightarrow 0$. To obtain a formal expression for the frequency spectrum of a crystal, let us first introduce the function $N(\omega^2)$ which gives the number of normal modes with squared frequencies less than or equal to ω^2 . A formal expression for $N(\omega^2)$ is

$$N(\omega^2) = \int_0^{\omega^2} \sum_j \delta(x - \omega_j^2) dx$$

(3.1.32)

since unity is added to the right side of this expression whenever the integration variable passes one of the normal mode frequencies. In terms of $N(\omega^2)$, the distribution function for the squares of the normal mode

frequencies $G(\omega^2)$ is given by

$$\begin{aligned} G(\omega^2) d\omega^2 &= \frac{1}{\mathcal{N}} \left[N(\omega^2 + d\omega^2) - N(\omega^2) \right] \\ &= \frac{1}{\mathcal{N}} N'(\omega^2) d\omega^2 + O((d\omega^2)^2) \end{aligned} \quad (3.1.33)$$

where \mathcal{N} ($= 3rN$) is the number of degrees of freedom in the crystal, so that

$$G(\omega^2) = \frac{1}{\mathcal{N}} \sum_j \delta(\omega^2 - \omega_j^2) \quad (3.1.34)$$

If we use the results that

$$\delta(x) = \frac{1}{\pi} \operatorname{Im} \lim_{\delta \rightarrow 0^+} \frac{1}{x - i\delta} \quad (3.1.35)$$

and that the frequency spectrum $g(\omega)$ is related to $G(\omega^2)$ by

$$g(\omega) = 2\omega G(\omega^2) \quad (3.1.36)$$

then we obtain the result

$$g(\omega) = \frac{2\omega}{\mathcal{N}} \frac{1}{\pi} \operatorname{Im} \lim_{\delta \rightarrow 0^+} \sum_j \frac{1}{\omega^2 - i\delta - \omega_j^2} \quad (3.1.37)$$

If we combine Eqs. (3.1.37) and (3.1.30) we see that

$$\operatorname{Im} \lim_{\delta \rightarrow 0^+} \frac{d}{d\omega} \ln |\Delta(\omega)| \Big|_{\omega^2 \rightarrow \omega^2 - i\delta}$$

$$\begin{aligned}
&= 2\omega \operatorname{Im} \lim_{\delta \rightarrow 0+} \sum_j \left\{ \frac{1}{\omega^2 - i\delta - \omega_j^2} - \frac{1}{\omega^2 - i\delta - \omega_{0j}^2} \right\} = \mathcal{N}\pi \\
&\times [g(\omega) - g_0(\omega)] = \mathcal{N}\pi \Delta g(\omega)
\end{aligned} \tag{3.1.38}$$

At this point we could well argue that knowing the change in the frequency spectrum of a crystal resulting from the introduction of defects, we have, at least formally, solved the problem of determining the change in the thermodynamic functions of the crystal. This is the point of view adopted by Lifshitz and his collaborators.¹ It should be remarked here that the calculation of the change in thermodynamic functions through the use of $\Delta g(\omega)$ is somewhat simplified if we first integrate the expression for the change by parts and work instead with the change in the integrated frequency spectrum, since the latter function usually has a simpler form than $\Delta g(\omega)$ itself. For example, we can write the change in the Helmholtz free energy as

$$\begin{aligned}
\Delta F &= \mathcal{N}kT \int_0^\infty \ln \left\{ 2 \sinh \frac{\hbar\omega}{2kT} \right\} \Delta g(\omega) d\omega \\
&= -\frac{\mathcal{N}\hbar}{2} \int_0^\infty \Delta n(\omega) \coth \frac{\hbar\omega}{2kT} d\omega
\end{aligned} \tag{3.1.39}$$

where

$$\Delta n(\omega) = \int_0^\omega \Delta g(\xi) d\xi \tag{3.1.40}$$

is the change in the integrated frequency spectrum. From Eqs. (3.1.40), (3.1.38), and (3.1.29) we find

$$\Delta n(\omega) = \frac{1}{\mathcal{N}\pi} \operatorname{Im} \lim_{\delta \rightarrow 0+} \ln |\Delta(\omega)| \omega^2 \rightarrow \omega^2 - i\delta \tag{3.1.41}$$

Equations (3.1.39) and (3.1.41) are Lifshitz' result for the change in free energy produced by defects.

However, from a purely practical point of view it is often more convenient, and certainly simpler, to compute changes in thermodynamic functions by methods which do not involve an explicit determination of $\Delta g(\omega)$ or $\Delta n(\omega)$. A similar situation exists in the computation of

the thermodynamic functions of perfect crystals, where a great deal of effort has gone into the development of techniques for their evaluation directly from the dynamical matrix of the crystal.¹⁶ The calculation of the change in the frequency spectrum resulting from defects is more difficult than the calculation of the spectrum of a perfect crystal. Very recently, a calculation of what is essentially $\Delta n(\omega)$ produced by a mass defect in a face-centered cubic crystal has been carried out by Nardelli and Tettamanzi.¹⁷ Their calculation suggests that the evaluation of $\Delta n(\omega)$ even for a mass defect in a crystal of lower symmetry, or for a more complicated defect in a cubic crystal, will present formidable (though not necessarily insurmountable) computational problems. In addition, it is often desirable, and even necessary, to know the analytic dependence of a change in an additive function of the normal mode frequencies on such parameters as the mass of an impurity atom and the force constants which characterize its interaction with the host crystal, and on the temperature. Such information is lost in a purely numerical calculation based on Eq. (3.1.39) or on Eq. (3.1.24) for that matter. For these reasons, although we keep in mind the possibility of using Eq. (3.1.39), in what follows we emphasize the development of techniques for evaluating the change in additive functions of the normal mode frequencies of a crystal as a result of defects, which do not require an explicit knowledge of either $\Delta g(\omega)$ or $\Delta n(\omega)$.

We now proceed to obtain expressions for the change in the Helmholtz free energy of a crystal resulting from the introduction of defects. From these expressions we can obtain the remaining thermodynamic functions.

Since we shall use the contour of Fig. III.1 in evaluating the integrals which appear in the theory, it is convenient to introduce a new function $\Omega(f)$ by the relation

$$\Omega(f) = \frac{d}{df} \ln |\Delta(i\omega_L f)| \quad (3.1.42a)$$

$$\Omega(f) = 2f \sum_j \left\{ \frac{1}{f^2 + f_j^2} - \frac{1}{f^2 + f_{0j}^2} \right\} \quad (3.1.42b)$$

The variable $f = \omega/\omega_L$ with similar definitions for f_j and f_{0j} .

To evaluate the change in the Helmholtz free energy by the methods of this chapter it is convenient to expand $f(\omega)$ as

$$f(\omega) = kT \ln \left\{ 2 \sinh \frac{\hbar \omega}{2kT} \right\} = \frac{1}{2} \hbar \omega - \frac{1}{\beta} \sum_{n=1}^{\infty} \frac{e^{-n\beta \hbar \omega}}{n} \quad (3.1.43)$$

where $\beta = (kT)^{-1}$. This form is valid for all temperatures but is

particularly useful at low temperatures. The expression for the change in the free energy becomes

$$\Delta F(T) = \Delta E_0 - kT \sum_{n=1}^{\infty} \frac{\Delta I_n(T)}{n} \quad (3.1.44)$$

where

$$\Delta E_0 = \frac{\hbar}{4\pi i} \int_c z \, d \ln | \Delta(z) | \quad (3.1.45)$$

$$\Delta I_n(T) = \frac{1}{2\pi i} \int_c e^{-n\beta \hbar z} \, d \ln | \Delta(z) | \quad (3.1.46)$$

If the contour of Fig. III. 1 is chosen, these expressions reduce to the following real integrals:

$$\Delta E_0 = -\frac{\hbar \omega_L}{2\pi} \int_0^{\infty} f \Omega(f) \, df \quad (3.1.47)$$

$$\Delta I_n(T) = \frac{1}{\pi} \int_0^{\infty} \Omega(f) \sin \alpha_n f \, df \quad (3.1.48)$$

where $\alpha_n = n\beta \hbar \omega_L$. At low temperatures α_n is very large and we can integrate Eq. (3.1.48) by parts to find

$$\Delta I_n(T) \sim \frac{1}{\pi} \left[\frac{\Omega(0)}{\alpha_n} - \frac{\Omega''(0)}{\alpha_n^3} + \frac{\Omega^{iv}(0)}{\alpha_n^5} - \dots \right] \quad (3.1.49)$$

If we substitute this expansion into Eq. (3.1.44) and carry out the sum over n we obtain finally

$$\begin{aligned} \Delta F(T) = \Delta E_0 - \frac{kT}{\pi} & \left\{ \frac{\pi^2}{6} \Omega(0) \left(\frac{kT}{\hbar \omega_L} \right) - \frac{\pi^4}{90} \Omega''(0) \left(\frac{kT}{\hbar \omega_L} \right)^3 \right. \\ & \left. + \frac{\pi^6}{945} \Omega^{iv}(0) \left(\frac{kT}{\hbar \omega_L} \right)^5 - \dots \right\} \end{aligned} \quad (3.1.50)$$

This expansion can be shown to be valid for one- and three-dimensional lattices [in the latter case $\Omega(0) = 0$]; however, for two-dimensional lattices $\Omega(f)$ has a branch point at $f = 0$, and the evaluation of the low-temperature expansion of $\Delta I_n(T)$ requires special techniques. This case is discussed in reference 18.

In the high-temperature limit the free energy can be expanded in inverse powers of the temperature

$$\begin{aligned} F(T) = kT \sum_j \ln \frac{\hbar \omega_j}{kT} + kT \sum_j \sum_{n=1}^{\infty} (-1)^{n+1} \\ \times \frac{B_{2n}}{2n(2n)!} \left(\frac{\hbar \omega_j}{kT} \right)^{2n} \end{aligned} \quad (3.1.51)$$

where the B_{2n} are the Bernoulli numbers

$$B_2 = \frac{1}{6} \quad B_4 = \frac{1}{30} \quad B_6 = \frac{1}{42} \dots \quad (3.1.52)$$

If we interchange the order of summations over j and n and recall the definition of the $2n$ -th moment of the frequency distribution function,

$$\mu_{2n} = \omega_L^{2n} u_{2n} = \frac{1}{N} \sum_j \omega_j^{2n} \quad (3.1.53)$$

we obtain

$$F(T) = kT \ln \prod_j \left(\frac{\hbar \omega_L}{kT} \right) f_j + N kT \sum_{n=1}^{\infty} (-1)^{n+1} \times \frac{B_{2n} u_{2n}}{(2n)! 2n} \left(\frac{\Theta}{T} \right)^{2n} \quad (3.1.54)$$

where we have put $\Theta = \hbar \omega_L / k$. This expansion is valid for all lattices and is independent of the dimensionality of the lattice. It converges for $(\Theta/T) < 2\pi$. However, methods exist for extending its validity to lower temperatures. These are presented in reference 19 and will not be discussed here.

We can rewrite the first term of Eq. (3.1.54) as

$$\frac{1}{2} kT \ln \prod_j \left(\frac{\hbar \omega_L}{kT} \right)^2 f_j^2 = N kT \ln \left(\frac{\hbar \omega_L}{kT} \right) + \frac{1}{2} kT \ln \prod_j f_j^2 \quad (3.1.55)$$

However, from Eq. (3.1.29) we see

$$\prod_j \frac{f_j^2}{f_{0j}^2} = \frac{A_0}{A} |\Delta(0)| \quad (3.1.56)$$

so that finally

$$\Delta F(T) = \frac{1}{2} k T \ln \frac{A_0}{A} |\Delta(0)| + \mathcal{N} k T \sum_{n=1}^{\infty} (-1)^{n+1} \times \frac{B_{2n} \Delta u_{2n}}{2n (2n)!} \left(\frac{\Theta}{T} \right)^{2n} \quad (3.1.57)$$

where Δu_{2n} is the change in the $2n$ -th moment resulting from defects.

An alternative derivation of Eq. (3.1.57) has been given by Yamahuzi and Tanaka.²⁰

Δu_{2n} can be calculated from $\Omega(f)$ in the following way. We expand $\Omega(f)$ for large $|f|$ in inverse powers of $|f|$, and from Eq. (3.1.42b) we obtain

$$\begin{aligned} \Omega(f) &= -\frac{2}{f} \sum_j \left\{ \frac{f_j^2 - f_{0j}^2}{f^2} - \frac{f_j^4 - f_{0j}^4}{f^4} + \frac{f_j^6 - f_{0j}^6}{f^6} - \dots \right\} \\ &= -\frac{2\mathcal{N}}{f} \sum_{n=1}^{\infty} (-1)^{n+1} \frac{\Delta u_{2n}}{f^{2n}} \end{aligned} \quad (3.1.58)$$

where we have used Eq. (3.1.53). Thus $\Omega(f)$ is a moment-difference generating function for an imperfect crystal.

It only remains to obtain an explicit expression for $|\Delta(u)|$.

In the presence of defects, the time independent equations of motion (3.1.8) take the form

$$M_{\kappa} \ddot{u}_{\alpha}^{(\iota)} - \sum_{\iota' \kappa' \beta} \mathfrak{F}_{\alpha \beta}^{(\iota \iota')} v_{\beta}^{(\iota')}$$

$$= \sum_{\ell' \kappa' \beta} \Delta_{\alpha \beta}(\ell \ell'; \kappa \kappa') v_{\beta}(\ell' \kappa') \quad (3.1.59)$$

where the matrix $\Delta_{\alpha \beta}(\ell \kappa; \ell' \kappa')$ describes the effects of the perturbation. If we define a Green's function $G_{\alpha \beta}(\ell \kappa; \ell' \kappa'; \omega)$ by the condition

$$\begin{aligned} \sum_{\ell'' \kappa'' \gamma} \left[M_{\kappa} \omega^2 \delta_{\ell \ell''} \delta_{\kappa \kappa''} \delta_{\alpha \gamma} - \Phi_{\alpha \gamma}(\ell \ell''; \kappa \kappa'') \right] G_{\alpha \beta}(\ell'' \kappa''; \ell' \kappa'; \omega) \\ = \delta_{\ell \ell'} \delta_{\kappa \kappa'} \delta_{\alpha \beta} \end{aligned} \quad (3.1.60)$$

then the formal solution to Eq. (3.1.59) is

$$v_{\alpha}(\ell' \kappa') = \sum_{\substack{\ell'' \kappa'' \beta \\ \ell'' \kappa'' \lambda}} G_{\alpha \beta}(\ell \ell'; \kappa \kappa'; \omega) \Delta_{\beta \lambda}(\ell' \ell''; \kappa' \kappa'') v_{\lambda}(\ell'' \kappa'') \quad (3.1.61)$$

Since Eq. (3.1.61) must hold for all $(\ell \kappa \alpha)$, it must be satisfied by the amplitudes $\{v_{\alpha}(\ell \kappa)\}$ which are those directly affected by the perturbation described by

$$\sum_{\ell' \kappa' \beta} \Delta_{\alpha \beta}(\ell \ell'; \kappa \kappa') v_{\beta}(\ell' \kappa')$$

Thus, if we set $(\ell \kappa \alpha)$ successively equal to the values of $(\ell'' \kappa'' \lambda)$ for which the right side of Eq. (3.1.61) is nonvanishing, we obtain a set of homogeneous equations for the displacement amplitudes directly affected by the defect. The determinantal equation which expresses the condition that this set of equations possess a nontrivial solution is the equation whose roots are the altered frequencies of the perturbed crystal.

We can obtain this result alternatively in the following way. If we identify the matrix $\Delta_{\alpha \beta}(\ell \kappa; \ell' \kappa')$ with the matrix $-\delta M$ of Eq. (3.1.23),

and the matrix $G_{\alpha\beta}(\ell\kappa; \ell'\kappa'; \omega)$ with $M_o^{-1}(\omega)$, then the determinant $|\Delta(w)^{\alpha\beta}|$ can be written explicitly

$$|\Delta(w)| = |\delta_{\ell\ell'} \delta_{\kappa\kappa'} \delta_{\alpha\beta} - \sum_{\ell''\kappa''\lambda} G_{\alpha\lambda}(\ell\ell''; \omega) \Delta_{\lambda\beta}(\ell''\ell'; \kappa''\kappa')| \quad (3.1.62)$$

The equation $|\Delta(w)| = 0$ is recognized to be nothing more than the condition that the set of homogeneous equations for the amplitudes directly involved in the perturbation have a nontrivial solution. The rank of the determinant $|\Delta(w)|$ is just the number of degrees of freedom in the crystal directly affected by the perturbation. For a localized defect this is a comparatively small number, and it is this fact that makes the methods described in this section so useful for treating defect problems.

To obtain an explicit expression for $G_{\alpha\beta}(\ell\kappa; \ell'\kappa'; \omega)$ we proceed essentially as in the one-dimensional case. We use the fact that

$$\delta_{\ell\ell'} \delta_{\kappa\kappa'} \delta_{\alpha\beta} = \frac{1}{N} \sum_{kj} e_{\alpha}(\kappa | \frac{k}{j}) e_{\beta}^{*}(\kappa' | \frac{k}{j}) \times e^{2\pi i k \cdot (x(\ell) - x(\ell'))} \quad (3.1.63)$$

and expand $G_{\alpha\beta}(\ell\kappa; \ell'\kappa'; \omega)$ as

$$G_{\alpha\beta}(\ell\ell'; \kappa\kappa'; \omega) = \frac{1}{N} \sum_{kj} f(\frac{k}{j}) \frac{e_{\alpha}(\kappa | \frac{k}{j})}{\sqrt{M_{\kappa}}} \frac{e_{\beta}^{*}(\kappa' | \frac{k}{j})}{\sqrt{M_{\kappa'}}} \times e^{2\pi i k \cdot (x(\ell) - x(\ell'))} \quad (3.1.64)$$

If we substitute Eq. (3.1.64) into Eq. (3.1.60) and use Eq. (3.1.63) we find that $f(k)$ is

$$f_j^{(k)} = \frac{1}{\omega^2 - \omega_j^2(k)} \quad (3.1.65)$$

with the result that

$$G_{\alpha\beta}(\ell\ell';\omega) = \frac{1}{N\sqrt{M_\alpha M_\beta}} \sum_{kj} \frac{e_{\alpha}(\alpha|k)_j e_{\beta}^{*}(\beta'|k)_j}{\omega^2 - \omega_j^2(k)} \times e^{2\pi i k \cdot (x(\ell) - x(\ell'))} \quad (3.1.66)$$

We conclude this section with a simple 1-D example of the application of the results we have so far obtained. We compute the change in the zero-point energy when the mass of the atom at the origin is changed from M to $M' = (1 - \epsilon) M$. From Eq. (2.2.14) we find that the determinant $|\Delta(\omega)|$ is

$$|\Delta(\omega)| = 1 - \epsilon M \omega^2 g_0(\omega) \quad (3.1.67)$$

from which we obtain the result

$$\Omega(f) = \frac{d}{df} \ln \left\{ 1 - \frac{2\epsilon f^2}{\pi} \int_0^{\frac{\pi}{2}} \frac{d\phi}{f^2 + \sin^2 \phi} \right\} \quad (3.1.68)$$

In obtaining this result we have replaced ω by $i\omega_L f$ in Eq. (2.2.7) and

have converted the sum into an integral, since the summand no longer has singularities at the unperturbed normal mode frequencies. We obtain explicitly

$$\Omega(f) = - \frac{\epsilon}{(1+f^2)^{3/2} - \epsilon f(1+f^2)} \quad (3.1.69)$$

According to Eq. (3.1.47), the change in the zero-point energy is given by

$$\begin{aligned} \Delta E_o &= \frac{\hbar \omega_L \epsilon}{2\pi} \int_0^{\pi} \frac{\sin \theta \, d\theta}{1 - \epsilon \sin \theta} \\ &= \frac{\hbar \omega_L}{2\pi} \left\{ -\frac{\pi}{2} + \frac{1}{\sqrt{1-\epsilon^2}} \cos^{-1}(-\epsilon) \right\} \end{aligned} \quad (3.1.70)$$

It should be emphasized that this is an exact result, not merely a small ϵ result. Since for $0 \leq \epsilon < 1$ the localized mode contributes

$$\Delta E_o^{\text{loc.}} = \frac{\hbar \omega_L}{2\pi} \left(\frac{\pi}{\sqrt{1-\epsilon^2}} - \pi \right) \quad 0 \leq \epsilon < 1 \quad (3.1.71)$$

to the zero-point energy (see Eq. 2.2.22), the in-band contribution is readily found to be

$$\Delta E_o^{\text{i.b.}} = \frac{\hbar \omega_L}{2\pi} \left[\frac{\pi}{2} + \frac{\cos^{-1}(-\epsilon) - \pi}{\sqrt{1-\epsilon^2}} \right] \quad (3.1.72)$$

and we see that the two contributions are of comparable magnitude. It also follows from Eq. (3.1.70) that ΔE_o is positive for $\epsilon > 0$ and negative for $\epsilon < 0$. These results are what would be expected in light of the general increase of all symmetric mode frequencies owing to the presence of a light defect and of their decrease owing to the presence of a heavy defect.

We now turn to an application of the results of this section to a three-dimensional defect problem.

IV. A MASS DEFECT IN A THREE-DIMENSIONAL CRYSTAL

1. Properties of the Normal Modes

To illustrate the rather formal results of the preceding Sec. III, we devote this section to a detailed study of the effects of a mass defect in a three-dimensional crystal. The results we obtain here are of interest in themselves and will also be useful in later discussions of dynamic properties of disordered crystals.

We shall now study the effects of the defect on the normal mode frequencies and vibration amplitudes. The change in the crystal's vibrational free energy produced by the defect will be discussed in Sec. IV. 2.

We consider first the case in which the impurity atom corresponds to the $\kappa = 1$ atom in the $\ell = 0$ unit cell. If we denote the mass of this atom by

$$M_1' = (1 - \epsilon) M_1 \quad (4.1.1)$$

the coefficient $\Delta_{\alpha\beta}(\ell\kappa; \ell'\kappa')$ takes the form

$$\Delta_{\alpha\beta}(\ell\kappa; \ell'\kappa') = \epsilon M_1 \omega^2 \delta_{\ell 0} \delta_{\ell' 0} \delta_{\kappa 1} \delta_{\kappa' 1} \delta_{\alpha\beta} \quad (4.1.2)$$

With this result the expression for the displacement components, Eq. (3.1.61), takes the form

$$v_{\alpha}(\ell\kappa) = \epsilon M_1 \omega^2 \sum_{\beta} G_{\alpha\beta}(\ell 0; \omega) v_{\beta}(\ell' 1) \quad (4.1.3)$$

If we put $\ell = 0$ and $\kappa = 1$, and let α assume the values x, y, z , successively, we obtain the eigenvalue equation for the normal mode frequencies of the perturbed crystal:

$$\begin{pmatrix} v_x^{(0)}(1) \\ v_y^{(0)}(1) \\ v_z^{(0)}(1) \end{pmatrix} = \epsilon M_1 \omega^2 \begin{pmatrix} G_{xx}^{(00)}(11; \omega) G_{xy}^{(00)}(11; \omega) G_{xz}^{(00)}(11; \omega) \\ G_{yx}^{(00)}(11; \omega) G_{yy}^{(00)}(11; \omega) G_{yz}^{(00)}(11; \omega) \\ G_{zx}^{(00)}(11; \omega) G_{zy}^{(00)}(11; \omega) G_{zz}^{(00)}(11; \omega) \end{pmatrix} \begin{pmatrix} v_x^{(0)}(1) \\ v_y^{(0)}(1) \\ v_z^{(0)}(1) \end{pmatrix} \quad (4.1.4)$$

The expression for $G_{\alpha\beta}(01; 01; \omega)$ is explicitly

$$G_{\alpha\beta}^{(00)}(11; \omega) = \frac{1}{NM_1} \sum_{kj} \frac{e_{\alpha}(1|k_j) e_{\beta}^*(1|k_j)}{\omega^2 - \omega_j^2(k)} \quad (4.1.5)$$

We now specialize to the case of a cubic crystal containing two ions per unit cell, each of which is at a center of inversion symmetry. The alkali halides provide an example of crystals of this type. In this special case the function $G_{\alpha\beta}(01; 01; \omega)$ can be simplified somewhat. The simplification stems from the fact that for such a crystal the product $e_{\alpha}(1|k_j) e_{\beta}^*(1|k_j)$ transforms as $k_{\alpha} k_{\beta}$ under any real orthogonal transformation of crystal axes which takes the crystal into itself. The frequency $\omega_j(k)$ is invariant against such a transformation. In the case of a crystal of the NaCl structure, the set of symmetry operations applied at a lattice point which take the crystal into itself coincide with the operations which take a face-centered cubic crystal into itself. These results have the consequence that $G_{\alpha\beta}(01; 01; \omega)$ is diagonal in α and β . It is moreover isotropic since the crystal is invariant under a rotation through 90° about any of the coordinate axes. We can thus write $G_{\alpha\beta}(01; 01; \omega)$ as

$$G_{\alpha\beta}^{(00)}(11; \omega) = \frac{\delta_{\alpha\beta}}{3NM_1} \sum_{kj} \frac{e(1|k_j) \cdot e^*(1|k_j)}{\omega^2 - \omega_j^2(k)} \quad (4.1.6)$$

The eigenvalue equation for the normal mode frequencies reduces to three separate equations, each of which is

$$1 = \epsilon \omega^2 \frac{1}{3N} \sum_{\mathbf{kj}} \frac{e(1 | \frac{\mathbf{k}}{j}) \cdot e^*(1 | \frac{\mathbf{k}}{j})}{\omega^2 - \omega_j^2(\mathbf{k})} \quad (4.1.7)$$

so that the frequencies of the perturbed crystal are triply degenerate.

We make no further use of the foregoing results here and put them aside for future reference.

We now turn our attention to Bravais lattices. The expression for the displacement components associated with a mass defect in such a crystal is obtained from Eq. (4.1.3) by suppressing the basis indices κ and κ' :

$$v_{\alpha}(t) = \epsilon M \omega^2 \sum_{\beta} G_{\alpha\beta}(t0; \omega) v_{\beta}(0) \quad (4.1.8)$$

The eigenvalue equation for the normal mode frequencies of the perturbed crystal becomes

$$\begin{pmatrix} v_x(0) \\ v_y(0) \\ v_z(0) \end{pmatrix} = \epsilon M \omega^2 \begin{pmatrix} G_{xx}(00; \omega) & G_{xy}(00; \omega) & G_{xz}(00; \omega) \\ G_{yx}(00; \omega) & G_{yy}(00; \omega) & G_{yz}(00; \omega) \\ G_{zx}(00; \omega) & G_{zy}(00; \omega) & G_{zz}(00; \omega) \end{pmatrix} \begin{pmatrix} v_x(0) \\ v_y(0) \\ v_z(0) \end{pmatrix} \quad (4.1.9)$$

For orthorhombic, tetragonal, and cubic crystals we can use symmetry arguments similar to those used above to show that the matrix with elements $G_{\alpha\beta}(00; \omega)$ is diagonal:

$$G_{\alpha\beta}(00;\omega) = \delta_{\alpha\beta} \frac{1}{NM} \sum_{kj} \frac{e_{\alpha}^{(k)} e_{\alpha}^{(k)}}{\omega^2 - \omega_j^2(k)} \quad (4.1.10)$$

This means that for orthorhombic crystals the eigenfrequencies are the roots of the three equations

$$1 = \epsilon \omega^2 \frac{1}{N} \sum_{kj} \frac{e_{\alpha}^{(k)} e_{\alpha}^{(k)}}{\omega^2 - \omega_j^2(k)} \quad \alpha = x, y, z \quad (4.1.11)$$

For tetragonal crystals with the fourfold rotation axis along the z-direction we obtain a pair of equations for the normal mode frequencies,

$$1 = \epsilon \omega^2 \frac{1}{N} \sum_{kj} \frac{e_x^{(k)} e_x^{(k)}}{\omega^2 - \omega_j^2(k)} \quad (\text{twice})$$

$$1 = \epsilon \omega^2 \frac{1}{N} \sum_{kj} \frac{e_z^{(k)} e_z^{(k)}}{\omega^2 - \omega_j^2(k)} \quad (4.1.12)$$

Finally, in the case of cubic crystals we obtain a single equation repeated three times

$$1 = \epsilon \omega^2 \frac{1}{3N} \sum_{kj} \frac{1}{\omega^2 - \omega_j^2(k)} \quad (4.1.13)$$

In all that follows we shall emphasize the case of cubic crystals, since they provide us with a practically important category of crystals, and one for which many of the calculations we shall encounter can be carried out analytically rather than numerically.

In the remainder of Sec. IV-1 we study the solutions of Eq. (4.1.13) and discuss their consequences for the displacement amplitudes.

In the one-dimensional example discussed in Sec. II, it was possible to evaluate the sum appearing on the right-hand side of Eq. (4.1.13) in closed form, so that the problem of obtaining solutions of this equation was greatly simplified. Unfortunately, no one has succeeded in effecting a similar exact evaluation of the three-dimensional form of the sum. However, an approximate way of solving Eq. (4.1.13) has been indicated by Lax.²¹

Let us replace ω_j^2 by λ , and let us denote the distinct eigenfrequencies of the unperturbed crystal $\{\omega_j^2(k)\}$ by $\{\lambda_j\}$, where $\lambda_1 < \lambda_2 < \lambda_3 < \dots < \lambda_L$, and λ_L is the square of the largest normal mode frequency. Then Eq. (4.1.13) can be written as

$$\frac{1}{\epsilon\lambda} = \frac{1}{3N} \sum_j \frac{z_j}{\lambda - \lambda_j} \quad (4.1.14)$$

where z_j is the multiplicity, or degeneracy, of the eigenfrequency λ_j and satisfies the normalization condition

$$\sum_j z_j = 3N \quad (4.1.15)$$

In both Eqs. (4.1.14) and (4.1.15) the sum over j is carried out over only the normal modes whose frequencies are distinct.

If we plot the right side of Eq. (4.1.14) against the left side we see that the solutions of this equation interlace the unperturbed normal mode frequencies $\{\lambda_j\}$: there is only one solution in the interval $(\lambda_r, \lambda_{r+1})$. This conclusion also follows from the fact that

$$\frac{d}{d\lambda} \left\{ \sum_j \frac{z_j}{\lambda - \lambda_j} \right\}$$

is always negative for λ in $(\lambda_r, \lambda_{r+1})$.

The separation between consecutive unperturbed eigenvalues is typically of order $L^{-1} = N^{-1/3}$, because this is the order of the separation between neighboring k -values. However, it must be remarked that this result is strictly applicable only in cases where the frequency spectrum has only a single branch, because in the case of overlapping branches, although the separation between neighboring eigenfrequencies may be $O(N^{-1/3})$ in each branch, it can happen that the separation between two eigenvalues belonging to two different branches is much smaller. In what follows, however, we make the assumption that $\lambda_j - \lambda_{j-1} = O(N^{-1/3})$.

We now ask for the shift in the p -th eigenfrequency, λ_p . Setting λ equal to $\lambda_p + \Delta\lambda_p$ we can rewrite Eq. (4.1.14)

$$\frac{1}{\epsilon(\lambda_p + \Delta\lambda_p)} = \frac{1}{3N} \frac{z_p}{\Delta\lambda_p} + \frac{1}{3N} \sum_j' \frac{z_j}{\lambda_p - \lambda_j + \Delta\lambda_p} \quad (4.1.16)$$

where the prime on the sum excludes the term with $j = p$. The summand in the second term of this equation is a smoothly varying function of j since each term is of no larger order than $O(N^{-2/3})$. This follows from the fact that $\lambda_j - \lambda_{j-1} = O(N^{-1/3})$ and $\Delta\lambda_p$ must be smaller than the distance between λ_{p+1} and λ_p , or between λ_p and λ_{p-1} depending on whether ϵ is positive or negative. We thus replace the sum in Eq. (4.1.16) by a Cauchy principle value integral over the frequency spectrum of the unperturbed crystal, and find that to lowest order in N^{-1}

$$\Delta\lambda_p = \frac{z_p}{3N} \epsilon\lambda_p \left\{ 1 - \epsilon\lambda_p \int_0^{\omega_L} \frac{g_0(\xi)}{(\lambda_p - \xi^2)_p} d\xi \right\}^{-1} \quad (4.1.17)$$

where $g_0(\xi)$ is the frequency distribution function of the perfect host crystal (normalized to unity), and ω_L is the maximum frequency of the perfect crystal. The frequency shift can be written in the more usual

notation as

$$\Delta \omega_j^2(k) = \frac{\epsilon z_j^{(k)}}{3N} \omega_j^2(k) \left[1 - \epsilon \omega_j^2(k) \int_0^L \frac{g_0(\xi) d\xi}{(\omega_j^2(k) - \xi^2)_p} \right]^{-1}$$

(4.1.18)

In view of the approximate nature of this result, which is the best anyone has succeeded in obtaining to date, it is certainly fortunate that we do not need to know the in-band frequency shifts in order to be able to calculate changes in additive functions of the normal mode frequencies.

In the case that ω is larger than the maximum frequency of the unperturbed host crystal we can rewrite Eq. (4.1.13) as

$$1 = \epsilon \omega^2 \int_0^L \frac{g_0(\xi)}{\omega^2 - \xi^2} d\xi$$

(4.1.19)

For ϵ in the interval $(0,1)$ we shall show that Eq. (4.1.19) can have a solution.

For all realistic models of crystals the function on the right side of Eq. (4.1.19) is finite and positive for $\omega > \omega_L$, and has its largest value for $\omega = \omega_L$. This means that unlike the situation in the one-dimensional case where a localized mode appears for any value of ϵ in the interval $(0,1)$, in the three-dimensional case ϵ has to exceed a certain critical value before a solution of Eq. (4.1.19) appears. This critical value of ϵ is given by

$$\epsilon_{cr} = \left\{ \omega_L^2 \int_0^L \frac{g_0(\xi) d\xi}{\omega_L^2 - \xi^2} \right\}^{-1}$$

(4.1.20)

For one- and two-dimensional crystals the integral in Eq. (4.1.20) diverges, so that here $\epsilon_{cr} = 0$. In this respect the Debye spectrum is also unrealistic, since the integral in Eq. (4.1.20) diverges if a Debye spectrum is employed.

Although the use of an electronic computer for the solution of Eq. (4.1.19) is a sufficient condition, it is not a necessary one. We proceed by expanding the right-hand side of Eq. (4.1.19) in inverse powers of ω^2 :

$$\begin{aligned} \frac{1}{\epsilon} &= \int_0^{\omega_L} g_0(\xi) \left[1 + \frac{\xi^2}{\omega^2} + \frac{\xi^4}{\omega^4} + \dots \right] d\xi \\ &= 1 + \frac{\mu_2}{\omega^2} + \frac{\mu_4}{\omega^4} + \dots \end{aligned} \quad (4.1.21)$$

where the equation

$$\mu_{2n} = \int_0^{\omega_L} \xi^{2n} g_0(\xi) d\xi \quad (4.1.22)$$

is an alternative definition of the $2n$ -th moment of the frequency spectrum of the unperturbed host crystal. Equation (4.1.21) can be put into a more convenient form if we introduce the dimensionless moments and frequency which are given by

$$u_{2n} = \frac{\mu_{2n}}{\omega_L^{2n}} \quad x = \frac{\omega}{\omega_L} \leq 1 \quad (4.1.23)$$

In terms of these quantities Eq. (4.1.21) becomes

$$\frac{1}{\epsilon} = 1 + u_2 x^2 + u_4 x^4 + u_6 x^6 + \dots \quad (4.1.24)$$

It is a general result that for all three-dimensional crystals whose

maximum frequency corresponds to an analytic or to a fluted maximum in the surfaces of constant frequency in reciprocal space, the frequency distribution function behaves like constant $(\omega_L^2 - \omega^2)^{1/2}$ in the limit as $\omega \rightarrow \omega_L$.²² It can be shown that such a behavior of the frequency spectrum leads to the following asymptotic form for the moments $\{u_{2n}\}$ in the limit of large n :²³

$$u_{2n} \sim \frac{A}{n^{3/2}} \quad (4.1.25)$$

The value of the constant A can be found if the expansion of $\omega_j^2(k)$ as a function of k about its maximum value is known. A cruder, but simpler, way of determining A is to fit Eq. (4.1.25) to the highest known moment.

The expansion (4.1.24) converges for $x = 1$, so that the critical value of ϵ which must be exceeded in order that a localized mode appear is given by

$$\frac{1}{\epsilon_{cr}} = 1 + u_2 + u_4 + u_6 + \dots \quad (4.1.26)$$

Moments of the frequency spectra of models of various different crystals have been computed by several authors. The most extensive calculations are those of Domb and Isenberg²⁴ who have computed the even moments up to μ_{30} for nearest and next nearest neighbor models of face-centered and body-centered cubic crystals for many different ratios of the next nearest to nearest neighbor force constants. We use their results for a nearest neighbor central force model of a face-centered cubic crystal to illustrate the dependence of the local mode frequency on the defect mass. The maximum frequency of this model is given by²⁵

$$\omega_L = \left(\frac{8 \phi''(r_0)}{M} \right)^{1/2} = \left(\frac{6 a_0 B}{M} \right)^{1/2} \quad (4.1.27)$$

where $\phi(r)$ is the interatomic potential, r_0 is the nearest neighbor separation, B is the bulk modulus, and $a_0 = r_0 \sqrt{2}$ is the lattice

parameter. The dimensionless moments for this model are²⁴

$$\begin{array}{lll}
 u_0 = 1.000000 & u_{10} = 0.138824 & u_{20} = 0.065643 \\
 u_2 = 0.500000 & u_{12} = 0.115788 & u_{22} = 0.058308 \\
 u_4 = 0.312500 & u_{14} = 0.098590 & u_{24} = 0.052123 \\
 u_6 = 0.222656 & u_{16} = 0.085196 & u_{26} = 0.046855 \\
 u_8 = 0.171631 & u_{18} = 0.074451 & u_{28} = 0.042332 \\
 & & u_{30} = 0.038421
 \end{array}
 \quad (4.1.28)$$

Fitting an expression of the form of Eq. (4.1.25) to u_{30} , we find that $A \approx 2.232$. This result, together with the fact that

$$\begin{aligned}
 \sum_{n=k}^{\infty} \frac{1}{n^{3/2}} &\sim \frac{2}{k^{1/2}} \left\{ 1 + \frac{1}{4k} + \frac{1}{16k^2} - \frac{7}{768k^4} \right. \\
 &\quad \left. + \frac{11}{2048k^6} + \dots \right\}
 \end{aligned}
 \quad (4.1.29)$$

enables us to evaluate the right-hand side of Eq. (4.1.26) with the result

$$\begin{aligned}
 \frac{1}{\epsilon_{cr}} &\approx \sum_{n=0}^{15} u_{2n} + \sum_{n=16}^{\infty} \frac{2.232}{n^{3/2}} \\
 &= 3.0233 + 1.1337 \\
 &= 4.1570
 \end{aligned}
 \quad (4.1.30)$$

This means that $\epsilon \geq \epsilon_{cr} = 0.2405$ in order that a localized mode can appear.

A table of ϵ as a function of x^2 calculated from Eq. (4.1.24) is given in Table IV.1.

We now ask, "What is the nature of the displacement amplitudes associated with in-band impurity modes and with modes whose frequencies have been removed out of the band?" For this purpose it is convenient to consider the behavior of the Green's function $G_{\alpha\beta}(\ell; \omega)$ for large $|x(\ell)|$, since it is this function which describes the displacements.

In the case of in-band modes we can write $G_{\alpha\beta}(\ell; \omega_j(k))$ as

$$\begin{aligned}
G_{\alpha\beta}(\mathbf{r}; \omega_j(\mathbf{k})) &= \frac{1}{NM} \sum_j^{z(\mathbf{k})} \frac{e_{\alpha}^{(\mathbf{k})} e_{\beta}^{(\mathbf{k})}}{\omega^2 - \omega_j^2(\mathbf{k})} e^{2\pi i \mathbf{k} \cdot \mathbf{x}(t)} \\
&+ \frac{1}{NM} \sum_{\mathbf{k}'_j} \frac{e_{\alpha}^{(\mathbf{k}'_j)} e_{\beta}^{(\mathbf{k}'_j)}}{(\omega_j^2(\mathbf{k}) - \omega_{j'}^2(\mathbf{k}'))_p} e^{2\pi i \mathbf{k}' \cdot \mathbf{x}(t)}
\end{aligned}
\tag{4.1.31}$$

Table IV. 1

The mass of an impurity atom as a function of the local mode frequency for a face-centered cubic crystal with nearest neighbor, central force interactions

x^2	ϵ	x	ω/ω_L	M'/M
0.0	1.0	0.0	∞	0
0.1	0.9494	0.3162	3.162	0.0506
0.2	0.8972	0.4472	2.236	0.1028
0.3	0.8432	0.5477	1.826	0.1568
0.4	0.7869	0.6325	1.581	0.2131
0.5	0.7277	0.7071	1.414	0.2723
0.6	0.6647	0.7746	1.291	0.3353
0.7	0.596	0.8367	1.195	0.404
0.8	0.519	0.8944	1.118	0.481
0.9	0.43	0.9487	1.054	0.57
1.0	0.2405	1.0000	1.000	0.7595

In the first term we replace $\omega^2 - \omega_j^2(\mathbf{k})$ by the expression given by Eq. (4.1.18). This results in a wavelike dependence of this term on $\mathbf{x}(\mathbf{r})$. In the second term we use the definition

$$\left(\frac{1}{x}\right)_p = \text{Im} \lim_{\delta \rightarrow 0^+} \int_0^{\infty} e^{ixt - \delta t} dt
\tag{4.1.32}$$

to rewrite it as

$$\frac{va}{M} \sum_{j'} \int_0^{\infty} dt \int_{Bz} d^3k' \sin(\omega_j^2(k) - \omega_{j'}^2(k')) t \\ \times e_{\alpha}^{(k')}(j') e_{\beta}^{(k')}(j') \cos 2\pi k' \cdot x(t) \quad (4.1.33)$$

where v_a is the volume of a unit cell of the crystal. An evaluation of this integral by the method of stationary phase for large $|x(t)|$ shows that it is wavelike as a function of $|x(t)|$. Thus the Green's function, and hence the displacement amplitudes, are wavelike in nature for those normal modes whose frequencies lie inside the band of allowed frequencies.

The displacement amplitudes associated with normal modes whose frequencies lie above the band of allowed frequencies decay faster than exponentially with increasing distance from the defect. This can be seen in the following way. We consider the contribution from a single branch only. The Green's function has the form

$$G_{\alpha\beta}^{(j)}(t; \omega) = \frac{va}{M} \int_{Bz} d^3k \frac{e_{\alpha}^{(k)}(j) e_{\beta}^{(k)}(j)}{\omega^2 - \omega_j^2(k)} e^{2\pi i k \cdot x(t)} \quad (4.1.34)$$

For large $|x(t)|$ the integrand oscillates rapidly as a function of k , and throughout most of the Brillouin zone contributions from different k -values tend to cancel. The exception to this comes from the region of k -values for which the denominator is as small as possible, i.e., the region about the maximum in the surfaces of constant frequency in k -space. About this point, which we denote by k_0 , $\omega_j^2(k)$ has the expansion

$$\omega_j^2(k) = \omega_L^2 - \frac{1}{2} \sum_{rs} A_{rs}^{(j)} (k_r - k_{0r}) (k_s - k_{0s}) + \dots$$

$$= \omega_L^2 - \frac{1}{2} \sum_{rs} A_{rs}^{(j)} \xi_r \xi_s + \dots \quad (4.1.35)$$

The linear terms vanish because we are expanding about a maximum. The matrix A is the matrix of a positive definite quadratic form. Assuming that there is only one maximum in the zone, we can now rewrite Eq. (4.1.34) as

$$G_{\alpha\beta}^{(j)}(t; \omega) \sim \frac{v_a}{M} e^{2\pi i k_0 \cdot x(t)} e_{\alpha}^{(k_0)} e_{\beta}^{(k_0)} \\ \times \int_0^\infty dt \int_{-\infty}^\infty d^3 \xi e^{2\pi i \sum_r \xi_r x_r(t) - t(\omega^2 - \omega_L^2 + \frac{1}{2} \sum_{rs} A_{rs}^{(j)} \xi_r \xi_s)} \quad (4.1.36)$$

where having made a small ξ expansion about k_0 we can then extend the ξ -integration over all space with little error. Since $e_{\alpha}(k_j)$ is ordinarily a smoothly varying function of k , we have replaced it by its value at $k = k_0$, and have taken it outside the integral. Strictly speaking, we should take into account the angular dependence of $e_{\alpha}(k_j)$ about the point $k = k_0$. This, however, will affect only the angular dependence of our result, and not its dependence on the magnitude of $x(t)$. In this way we obtain the result

$$G_{\alpha\beta}^{(j)}(t; \omega) \sim e^{2\pi i k_0 \cdot x(t)} \frac{v_a}{M} e_{\alpha}^{(k_0)} e_{\beta}^{(k_0)} \\ \times \int_0^\infty dt e^{-t(\omega^2 - \omega_L^2)}$$

$$\begin{aligned}
& \times \frac{(2\pi)^{3/2}}{|A|^{1/2}} \frac{1}{t^{3/2}} e^{-\frac{2\pi^2}{t}} \sum_{rs} A_{rs}^{-1} x_r(t) x_s(t) \\
& = e^{2\pi i k_o \cdot x(t)} \frac{v_a}{M} e_{\alpha}(k_o) e_{\beta}(k_o) \frac{(2\pi)}{|A|^{1/2}} \\
& \times \frac{\left\{ -2\sqrt{\omega^2 - \omega_L^2} \sqrt{2\pi} \left(\sum_{rs} A_{rs}^{-1} x_r(t) x_s(t) \right)^{1/2} \right\}}{\sqrt{\sum_{rs} A_{rs}^{-1} x_r(t) x_s(t)}}
\end{aligned}
\tag{4.1.37}$$

We see that the rate of exponential decay increases the higher ω is above the top of the allowed band.

It can occur that because the three branches of the frequency spectrum have different cut-off frequencies, a frequency can be out of the band for one or two of the branches and still in the band for the third branch. In such a case we see from the foregoing analysis that the displacements at large distances will be determined only by the in-band modes, with the out of band contributions being exponentially small.

2. Thermodynamic Functions of a Cubic Bravais Crystal Containing a Single Mass Defect

In the case that a mass defect is introduced into a cubic Bravais crystal the determinant $|\Delta(\omega)|$ obtained from Eq. (4.1.9) is

$$|\Delta(\omega)| = [1 - \epsilon M \omega^2 G_{xx}(0; \omega)]^3 \tag{4.2.1}$$

Consequently the function $\Omega(f)$ is given by

$$\Omega(f) = \frac{d}{df} \ln |\Delta(i\omega_L f)| = \frac{3d}{df} \ln \left\{ 1 - \frac{\epsilon \omega_L^2 f^2}{3N} \right. \\ \left. \times \sum_{\substack{k \\ j}} \frac{1}{\omega_L^2 f^2 + \omega_j^2(k)} \right\} \quad (4.2.2)$$

Since the summand in this expression is nonsingular, we can pass to the limit $N \rightarrow \infty$ and replace summation over k and j by integration over the frequency spectrum. In this way we obtain

$$\Omega(f) = -3\epsilon \frac{2f G(f) + f^2 G'(f)}{1 - \epsilon f^2 G(f)} \quad (4.2.3)$$

where we have introduced the functions

$$G(f) = \int_0^1 \frac{P(x)}{f^2 + x^2} dx \\ G'(f) = -2f \int_0^1 \frac{P(x) dx}{(f^2 + x^2)^2} \quad (4.2.4)$$

which are defined in terms of the dimensionless frequency spectrum,

$$P(x) = \omega_L g_0(\omega_L x) \quad (4.2.5)$$

The change in zero-point energy is given by

$$\Delta E_0 = \frac{3\epsilon \hbar \omega_L}{2\pi} \int_0^\infty \frac{2f^2 G(f) + f^3 G'(f)}{1 - \epsilon f^2 G(f)} df$$

(4. 2. 6)

The frequency spectrum $P(x)$ is known numerically for many models of different cubic crystals. The functions $G(f)$ and $G'(f)$ can then be obtained by numerical integration, and a second numerical integration yields ΔE_0 . Essentially this procedure has been carried out in connection with a closely related problem which will be discussed in Sec. IV-3. However, since our aim in this section is to illustrate techniques as well as to obtain specific results, we resort here to a simpler procedure. If we assume a Debye frequency spectrum for the host crystal

$$g_0(\omega) = \frac{3\omega^2}{\omega_D^3} \quad 0 < \omega < \omega_D$$

$$= 0 \quad \omega_D < \omega$$

we find

$$G(f) = 3 \left(1 - f \tan^{-1} \frac{1}{f} \right) \quad (4. 2. 7a)$$

$$G'(f) = 3 \left(\frac{f}{1+f^2} - \tan^{-1} \frac{1}{f} \right) \quad (4. 2. 7b)$$

The change in zero-point energy is then given by

$$\Delta E_0 = \frac{9\epsilon \hbar \omega_D}{2\pi} \int_0^\infty \frac{2f^2 - 3f^3 \tan^{-1} \frac{1}{f} + \frac{f^4}{1+f^2}}{1 - 3\epsilon (f^2 - f^3 \tan^{-1} \frac{1}{f})} df$$

$$= \frac{9\epsilon \hbar \omega_D}{2\pi} \int_0^{\frac{\pi}{2}} \frac{2 \sin x \cos^2 x + \sin x \cos^4 x - 3x \cos^3 x}{\sin^3 x - 3\epsilon (\sin x \cos^2 x - \cos^3 x)} \frac{dx}{\sin^2 x}$$

(4. 2. 8)

This integral has been evaluated numerically for several values of ϵ and the results are given in Table IV. 2.

Table IV. 2

The change in the zero-point energy of a crystal containing a mass defect, calculated in the Debye approximation, as a function of ϵ . The change ΔE_0 is given by

$$(9\hbar\omega_D/2\pi) \epsilon I(\epsilon)$$

ϵ	$I(\epsilon)$
-4	0.03434 π
-2	0.05261 π
-1	0.07306 π
1/4	0.1555 π
1/2	0.2089 π
3/4	0.3378 π

To calculate the change in the Helmholtz free energy at low but finite temperatures, according to Eq. (3.1.50) we must know the small $|f|$ behavior of $\hat{\rho}(f)$. If we make use of the fact that at low frequencies the frequency spectrum of a three-dimensional crystal has the expansion

$$g_0(\omega) = A_2\omega^2 + A_4\omega^4 + A_6\omega^6 + \dots$$

(4. 2. 9)

then it can be shown that the small $|f|$ expansion of $G(f)$ is given by²⁶

$$G(f) \sim \omega_L^2 \mu_{-2} - \frac{\pi}{2} A_2 \omega_L^3 |f| + C_2 f^2 + \frac{\pi}{2} A_4 \omega_L^5 |f|^3 + C_4 f^4 + \dots$$

(4. 2. 10)

where C_2, C_4, \dots , are coefficients which cannot be expressed simply in terms of the $\{A_{2n}\}$. It is then straightforward to obtain the results that

$$\begin{aligned}\Omega(0+) &= 0 \\ \Omega''(0+) &= 9\epsilon\pi\omega_L^3 A_2 \\ \Omega^{iv}(0+) &= 180\epsilon\pi\omega_L^5 (\epsilon A_2 \mu_{-2} - A_4)\end{aligned}\tag{4.2.11}$$

According to Eq. (3.1.50) the low-temperature expansion for the change in the free energy is given by

$$\begin{aligned}\Delta F(T) \sim \Delta E_0 + \frac{\hbar\omega_L}{\pi} \left\{ \frac{\pi^5}{10} \epsilon\omega_L^3 A_2 \left(\frac{kT}{\hbar\omega_L} \right)^4 \right. \\ \left. - \frac{4\pi^7}{21} \epsilon\omega_L^5 (\epsilon A_2 \mu_{-2} - A_4) \left(\frac{kT}{\hbar\omega_L} \right)^6 + O(T^8) \right\}\end{aligned}\tag{4.2.12}$$

Experimental values of the minus second moment μ_{-2} can be obtained from the temperature dependence of the Debye-Waller factor at high temperatures (see Sec. IX). The coefficient A_2 is related to the 0°K value of the equivalent Debye characteristic temperature Θ_0 by

$$A_2 = 3 \left(\frac{\hbar}{k\Theta_0} \right)^3\tag{4.2.13}$$

Finally, values of the coefficient A_4 have been calculated for models of several face-centered cubic crystals by Horton and Schiff²⁷ and by Marcus.²⁸

In the high-temperature limit we need the large $|f|$ expansion of $\Omega(f)$. We readily find

$$G(f) \sim \frac{1}{f^2} - \frac{u_2}{f^4} + \frac{u_4}{f^6} - \frac{u_6}{f^8} + \dots$$

(4.2.14)

where the u_{2n} are the dimensionless moments

$$u_{2n} = \int_0^1 x^{2n} P(x) dx$$

(4.2.15)

With the aid of this result we obtain the expansion

$$\Omega(f) \sim \frac{2}{f} (3N) \left\{ -\frac{1}{3N} \frac{3\epsilon}{1-\epsilon} \frac{u_2}{f^2} + \frac{1}{3N} \frac{6\epsilon}{1-\epsilon} \frac{u_4}{f^4} + \frac{1}{3N} \frac{3\epsilon^2}{(1-\epsilon)^2} \frac{u_2^2}{f^4} + O(f^{-6}) \right\}$$

(4.2.16)

from which, together with Eq. (3.1.58), we obtain

$$\Delta u_2 = \frac{1}{3N} \frac{3\epsilon}{1-\epsilon} u_2$$

$$\Delta u_4 = \frac{1}{3N} \left[\frac{6\epsilon}{1-\epsilon} u_4 + \frac{3\epsilon^2}{(1-\epsilon)^2} u_2^2 \right]$$

(4.2.17)

Then using the result that $|A(0)| = 1$, we find that the change in the Helmholtz free energy at high temperatures is given by

$$\Delta F(T) = -\frac{3kT}{2} \ln(1-\epsilon) + kT \left\{ \frac{1}{24} \left(\frac{\hbar \omega_L}{kT} \right)^2 \frac{3\epsilon}{1-\epsilon} u_2 \right.$$

$$- \frac{1}{2880} \left(\frac{\hbar \omega_L}{kT} \right)^4 \left[\frac{6\epsilon}{1-\epsilon} u_4 + \frac{3\epsilon^2}{(1-\epsilon)^2} u_2^2 \right] + O(T^{-6}) \Big\} \quad (4.2.18)$$

The first term inside the curly brackets can be formally simplified if we introduce the limiting high-temperature value of the equivalent Debye temperature²⁹

$$\Theta_\infty = \frac{\hbar}{k} \sqrt{\frac{5}{3}} \mu_2 \quad (4.2.19)$$

If anharmonic effects are small, it is possible to obtain Θ_∞ from experimental specific heat data. We accordingly obtain

$$\begin{aligned} \Delta F(T) = & - \frac{3kT}{2} \ln(1-\epsilon) + kT \left\{ \frac{3}{40} \frac{\epsilon}{1-\epsilon} \left(\frac{\Theta_\infty}{T} \right)^2 \right. \\ & \left. - \frac{1}{960} \left(\frac{\hbar \omega_L}{kT} \right)^4 \left[\frac{2\epsilon}{1-\epsilon} u_4 + \frac{\epsilon^2}{(1-\epsilon)^2} u_2^2 \right] + \dots \right\} \end{aligned} \quad (4.2.20)$$

This completes the catalogue of the effect on the free energy of a cubic Bravais crystal resulting from a single mass defect. We see that qualitatively the effect of adding a light impurity atom to the crystal ($0 < \epsilon < 1$) is to increase the free energy, whereas the addition of a heavy impurity ($\epsilon < 0$) reduces the free energy. This is to be expected since $kT \ln \{ 2 \sinh \hbar \omega / 2kT \}$ is a monotonically increasing function of ω for fixed T . We wish to emphasize again that none of these results are small ϵ results but are valid for all $\epsilon < 1$. This fact illustrates the advantages of using the Green's function formalism.

If we have a very dilute solution of mass defects in a cubic Bravais crystal, then it has been shown in the paper by Montroll, Maradudin, and Weiss listed in reference 2 that the change in the free energy of the host crystal is given correctly to terms linear in the concentration of impurities simply by multiplying the results of this section by the impurity concentration. In Sec. V this result will be obtained in a different manner.

3. Example: The Second-Order Doppler Shift in the Energy of a Recoilless Gamma-Ray Photon

The preceding discussion has dealt with studies of the effects of a defect on individual vibration modes of a crystal. In this section I wish to discuss a problem in which the dynamical behavior of an individual impurity atom is of primary interest.

Whenever a nucleus bound in a crystal decays from an excited state to its ground state by recoilless γ emission, the energy of the emitted γ ray is not just that given by the relativistic expression $E = \delta m c^2$, where δm is the difference in the nuclear mass between its ground and excited states. The decrease in nuclear mass during the γ -ray emission increases the normal mode frequencies of the crystal, in accordance with Rayleigh's theorems; and hence the energy of the crystal increases slightly. This energy must come from somewhere and in fact is drained off from the γ -ray. A very simple derivation of the expression for the fractional shift in the γ -ray energy produced by this process has been given by Josephson,³⁰ and proceeds as follows.

If the i -th nucleus in the crystal emits a γ -ray of energy E its mass changes by the amount $\delta m_i = -E/c^2$. Because of this change in mass, the energy of the crystal changes by the amount of the change in the mean kinetic energy of that nucleus. This change is given by

$$\delta E = \langle \Delta H \rangle = -\delta m_i \left\langle \frac{p_i^2}{2m_i} \right\rangle = \frac{E}{m_i c^2} \left\langle \frac{p_i^2}{2m_i} \right\rangle$$

(4.3.1)

and thus requires a knowledge of the mean kinetic energy of the emitting atom. Usually the emitting atom is an impurity in a host crystal so that its mean kinetic energy is somewhat different from that of one of the atoms in the pure host crystal. It is to the calculation of the mean kinetic energy of an impurity atom in a crystal that we now turn.

The interest in such a calculation stems from the fact that, when one attempts to measure the above-mentioned shift in the γ -ray energy, called the "second-order Doppler shift," by a Mössbauer-type experiment,³¹ in general one measures not the second-order Doppler shift alone but an additional shift, called the "isomeric shift," as well.^{31a} This total shift I shall call the "center shift." The isomeric shift arises essentially from the difference between the radius of a nucleus in its excited state and in its ground state and depends on the electronic charge density at the nucleus. It is consequently of some interest to nuclear physicists. In

order that the value of the isomeric shift can be obtained from an experimental determination of the center shift, we have to be able to subtract off the value of the shift given by Eq. (4.3.1). The isomeric shift can be taken to be independent of temperature, and this fact ultimately provides us with a means for accomplishing this subtraction.

The expectation value of an operator O in the canonical ensemble is given by

$$\langle O \rangle = \frac{\text{Tr } e^{-\beta H_O}}{\text{Tr } e^{-\beta H}} \quad (4.3.2)$$

where H is the Hamiltonian for the system described by the ensemble.

In the present case our system consists of a crystal, one of whose atoms is replaced by an impurity atom whose mass differs from that of the atom it replaces and whose potential of interaction with the other atoms in the crystal differs from that of the normal atom at that site. If we denote the mass of the κ -th atom in the ℓ -th unit cell by M_{κ}^{ℓ} , and if we assume that the impurity atom is the $\kappa = 1$ atom in the $\ell = 0$ unit cell, the Hamiltonian for the perturbed crystal can be written

$$\begin{aligned} H &= \sum_{\ell, \kappa} \frac{p_{\kappa}^{2(\ell)}}{2M_{\kappa}^{\ell}} + V \\ &= \left(\frac{1}{2M_1^{(0)}} - \frac{1}{2M_1} \right) p_1^{2(0)} + \sum_{\ell, \kappa} \frac{p_{\kappa}^{2(\ell)}}{2M_{\kappa}^{\ell}} + V \end{aligned} \quad (4.3.3)$$

In Eq. (4.3.3) M_{κ} is the mass of the κ -th kind of atom in the unit cell of the perturbed crystal, and we have omitted the now unnecessary superscript ℓ on the masses in the second sum of this equation. V represents the potential energy of the crystal and contains the effects of any force constant changes associated with the introduction of the impurity. If we write the mass of the impurity atom as

$$M_1^{(0)} = (1 - \epsilon) M_1 \quad (4.3.4)$$

the kinetic energy operator of the crystal can be written

$$T = \frac{\epsilon}{1-\epsilon} \frac{p_{(1)}^{2(0)}}{2M_1} + \sum_{\ell, \kappa} \frac{p_{(\ell, \kappa)}^{2(\ell)}}{2M_{\kappa}} \quad (4.3.5)$$

The operator O in our case is

$$O = \frac{p_{(1)}^{2(0)}}{2M_1^{(0)}} = \frac{1}{1-\epsilon} \frac{p_{(1)}^{2(0)}}{2M_1} \quad (4.3.6)$$

so that Eq. (4.3.2) becomes

$$\begin{aligned} \frac{\langle p_{(1)}^{2(0)} \rangle}{2M_1^{(0)}} &= \frac{\text{Tr } e^{-\beta \frac{\epsilon}{1-\epsilon} \frac{p_{(1)}^{2(0)}}{2M_1} - \beta H_0} \frac{p_{(1)}^{2(0)}}{2M_1^{(0)}}}{\text{Tr } e^{-\beta \frac{\epsilon}{1-\epsilon} \frac{p_{(1)}^{2(0)}}{2M_1} - \beta H_0}} \\ &= -\frac{1}{\beta(1-\epsilon)} \frac{d}{d(\frac{\epsilon}{1-\epsilon})} \ln \text{Tr } e^{-\beta \frac{\epsilon}{1-\epsilon} \frac{p_{(1)}^{2(0)}}{2M_1} - \beta H_0} \\ &= (1-\epsilon) \frac{d}{d\epsilon} F(\epsilon, T) \end{aligned} \quad (4.3.7)$$

where $F(\epsilon, T)$ is the Helmholtz free energy of the perturbed crystal. $F(\epsilon, T)$ can be written as the sum of two contributions,

$$F(\epsilon, T) = F_0(T) + \Delta F(\epsilon, T) \quad (4.3.8)$$

where $F_0(T)$ is the free energy of the unperturbed crystal, and

$\Delta F(\epsilon, T)$ is the change in the free energy resulting from the introduction of the impurity. $\Delta F(\epsilon, T)$ depends both on the parameter ϵ that characterizes the mass change and on the parameters that characterize the force constant changes as well, although we have not indicated the latter dependence explicitly. We have thus obtained the result that the mean kinetic energy of the impurity atom is given by

$$\frac{\langle p^2(1) \rangle}{2M_1(0)} = (1 - \epsilon) \frac{d}{d\epsilon} \Delta F(\epsilon, T) \quad (4.3.9)$$

The second-order Doppler shift in the γ -ray energy is, accordingly, given by

$$\frac{\delta E}{E} = \frac{1}{M_1 c^2} \frac{d}{d\epsilon} \Delta F(\epsilon, T) \quad (4.3.10)$$

This result is convenient since we can use the results of Sec. III for the change in the Helmholtz free energy resulting from defects to obtain δE directly.

Thus, for example, if the emitting nucleus can be approximated by a mass defect, we find from the results of the preceding section that the relative shift in a cubic Bravais host crystal is given by

$$\left(\frac{\delta E}{E} \right)_{0^\circ K} = \frac{1}{Mc^2} \frac{3\hbar\omega_L}{2\pi} \int_0^\infty \frac{2f^2 G(f) + f^3 G'(f)}{[1 - \epsilon f^2 G(f)]^2} df \quad (4.3.11a)$$

$$\begin{aligned} \left(\frac{\delta E}{E} \right)_{T=0} &= \left(\frac{\delta E}{E} \right)_{0^\circ K} + \frac{1}{Mc^2} \frac{\hbar\omega_L}{\pi} \left\{ \frac{\pi^5}{10} A_2 \omega_L^3 \left(\frac{kT}{\hbar\omega_L} \right)^4 \right. \\ &\quad \left. - \frac{4\pi^7}{21} \omega_L^5 (2\epsilon A_2 \mu_{-2} - A_4) \left(\frac{kT}{\hbar\omega_L} \right)^6 + \dots \right\} \end{aligned} \quad (4.3.12)$$

$$\left(\frac{\delta E}{E}\right)_{T \rightarrow \infty} = \frac{3kT}{2M'c^2} \left\{ 1 + \frac{1}{20} \frac{M}{M'} \left(\frac{\Theta_{\infty}}{T}\right)^2 - \frac{1}{720} \left(\frac{\hbar\omega_L}{kT}\right)^4 \left[\frac{u_4}{1-\epsilon} + \frac{\epsilon u_2^2}{(1-\epsilon)^2} \right] + \dots \right\} \quad (4.3.13)$$

The zero temperature shift can be evaluated numerically as a function of ϵ once the frequency spectrum of the host crystal is known. If we assume a Debye spectrum, we find that

$$\left(\frac{\delta E}{E}\right)_{0^\circ K} = \frac{1}{Mc^2} \frac{9\hbar\omega_D}{2\pi} \times \int_0^{\frac{\pi}{2}} \frac{2\sin^2 x \cos^2 x + \sin^2 x \cos^4 x - 3x \cos^3 x \sin x}{[\sin^3 x - 3\epsilon(\sin x \cos^2 x - x \cos^3 x)]^2} dx \quad (4.3.11b)$$

The integral in Eqs. (4.3.11a) has been evaluated numerically as a function of ϵ for a moderately realistic model of copper as the host crystal.²⁶ The integral in Eq. (4.3.11b) was also evaluated numerically assuming a value for ω_D appropriate to copper.²⁶ The results for the mean kinetic energy of the impurity atom are plotted in Fig. IV.1. The results suggest that the use of the Debye spectrum is not too bad an approximation for light impurities, but becomes a poorer approximation for heavy impurities.

The preceding results have been obtained for the case of a substitutional impurity whose mass is different from that of the atom it replaces. Of equal interest is the defect whose potential of interaction with the atoms of the host crystal, as well as its mass, is different from that of the atom it replaces. At the present time it appears to be possible to evaluate the second-order Doppler shift resulting from such an impurity without heavy numerical work only in the high-temperature limit. In this case we have

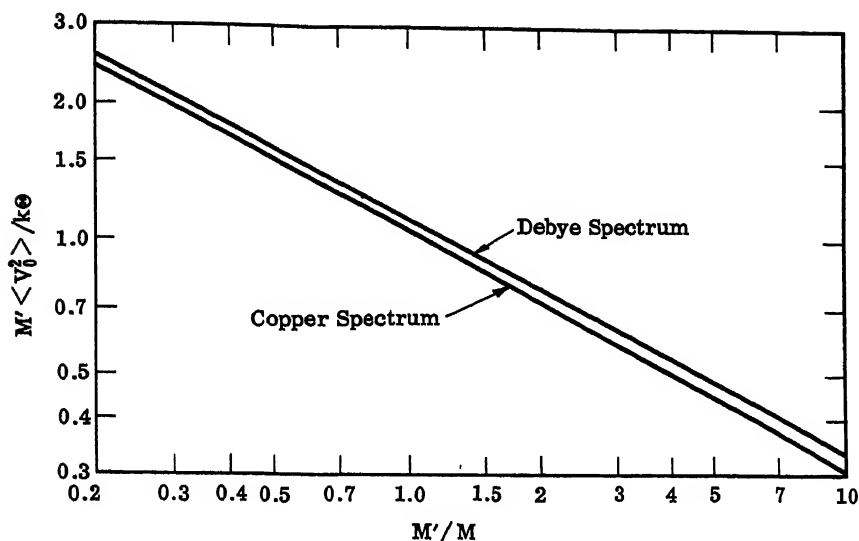


Fig. IV.1 The mean kinetic energy at 0°K of a mass defect atom in copper plotted as a function of the ratio of the impurity mass to the mass of a copper atom. Θ is the 0°K value of the Debye characteristic temperature for copper. One curve is calculated on the basis of a realistic frequency spectrum for copper, and the second is based on a Debye spectrum for copper.

$$\left(\frac{\delta E}{E} \right)_{T \rightarrow \infty} = \frac{1}{Mc^2} \left\{ \frac{3kT}{2} \frac{1}{1-\epsilon} + \frac{kT}{24} \left(\frac{\hbar}{kT} \right)^2 \right. \\ \left. \times \frac{d}{d\epsilon} (3N \Delta \mu_2) + \dots \right\} \quad (4.3.14)$$

The first term of this expansion does not reflect the change in the atomic force constants since it is a general result that the mean kinetic energy of an atom in the classical limit ($\hbar \rightarrow 0$) is independent of the potential energy.²⁶ The leading correction to the classical result, however, already shows the effects of the force constant changes, and these can be determined as follows. For simplicity, we consider only a Bravais

crystal; however, the generalization to nonprimitive crystals is straightforward.

We require the change in the second moment of the frequency spectrum of the host crystal, resulting from the introduction of the impurity. Let us rewrite Eq. (3.1.55) in the form

$$M_l \omega^2 v_\alpha(l) - \sum_{l', \beta} \left\{ \Phi_{\alpha\beta}(ll') + \Delta \Phi_{\alpha\beta}(ll') \right\} v_\beta(l') = 0$$

(4.3.15)

where $\Delta \Phi_{\alpha\beta}(ll')$ describes the change in the force constant matrix produced by the impurity, and M_l is the mass of the l -th atom. If we replace the $\{v_\alpha(l)\}$ by new amplitudes

$$v_\alpha(l) = \frac{\dot{v}_\alpha(l)}{\sqrt{M_l}}$$

we see that the normal mode frequencies are the eigenvalues of the matrix whose elements are

$$D_{\alpha\beta}(ll') = \frac{\Phi_{\alpha\beta}(ll')}{\sqrt{M_l M_{l'}}} + \frac{\Delta \Phi_{\alpha\beta}(ll')}{\sqrt{M_l M_{l'}}}$$

(4.3.16)

Since the sum of the n -th powers of the eigenvalues of a matrix equals the trace of the n -th power of the matrix, we find that the second moment of the frequency spectrum of the perturbed crystal is given by

$$3N\mu_2 = \sum_{l\alpha} \frac{\Phi_{\alpha\alpha}(ll) + \Delta \Phi_{\alpha\alpha}(ll)}{M_l}$$

(4.3.17)

If we assume that the impurity atom occupies the cell $\ell = 0$, then we can write Eq. (4.3.17) in the form

$$3N \Delta \mu_2 = \frac{\epsilon}{1 - \epsilon} \frac{1}{M} \sum_{\alpha} [\bar{\Phi}_{\alpha\alpha}(0) + \Delta \bar{\Phi}_{\alpha\alpha}(0)]$$

$$+ \frac{1}{M} \sum_{\ell\alpha} \Delta \bar{\Phi}_{\alpha\alpha}(\ell)$$
(4.3.18)

where we have used the fact that $\bar{\Phi}_{\alpha\beta}(\ell\ell')$ depends on ℓ and ℓ' only through their difference. It follows that

$$\frac{d}{d\epsilon} (3N \Delta \mu_2)$$

$$= \frac{1}{M (1 - \epsilon)^2} \sum_{\alpha} [\bar{\Phi}_{\alpha\alpha}(0) + \Delta \bar{\Phi}_{\alpha\alpha}(0)]$$
(4.3.19)

It is a general result, however, that

$$\bar{\Phi}_{\alpha\alpha}(0) = - \sum'_{\ell} \bar{\Phi}_{\alpha\alpha}(\ell)$$
(4.3.20a)

$$\Delta \bar{\Phi}_{\alpha\alpha}(0) = - \sum'_{\ell} \Delta \bar{\Phi}_{\alpha\alpha}(\ell)$$
(4.3.20b)

where the prime on the sums indicates that the point $\ell = 0$ is to be omitted. Equations (4.3.20) follow from the fact that an arbitrary, rigid

body translation of the crystal cannot change the net force on an atom in the equilibrium configuration.

By combining Eqs. (4.3.14), (4.3.19), and (4.3.20), we obtain finally

$$\left(\frac{\delta E}{E} \right)_{T \rightarrow \infty} = \frac{3kT}{2M'c^2} \left\{ 1 - \frac{1}{36} \left(\frac{\hbar}{kT} \right)^2 \frac{1}{M(1-\epsilon)} \right. \\ \left. \times \sum_{\alpha} [\bar{\Phi}_{\alpha\alpha}(\iota) + \Delta \bar{\Phi}_{\alpha\alpha}(\iota)] + O(T^{-4}) \right\} \quad (4.3.21)$$

This result holds for an arbitrary Bravais lattice. Since $\bar{\Phi}_{\alpha\alpha}(\iota)$ is assumed to be known, as is ϵ , in a given experiment, measurements of $(\delta E/E)$ at high temperatures can yield some information about $\Delta \bar{\Phi}_{\alpha\alpha}(\iota)$. The contribution from higher moments can be obtained in a similar fashion, but rapidly become more cumbersome.

I don't want to go into a detailed discussion of how an experimental measurement of the center shift is carried out. However, I should like to say enough about this point to show how the preceding theoretical results can be used together with experimental values for the center shift to obtain information about the isomeric shift.

Experimentally, γ -ray photons are generated by the recoilless transition from an excited state to its ground state of an appropriate nucleus bound in one crystal (the source). The γ ray is then absorbed by a nucleus of the same kind which is present in a different crystal (the absorber). The absorbing nucleus is raised to an excited state. The energy of the emitted γ ray can be written as $E + \Delta E_S$ (isomeric) + ΔE_S (s. o. D.), whereas the energy change of the absorbing nucleus is $E + \Delta E_A$ (isomeric) + ΔE_A (s. o. D.). The subscripts S and A refer to source and absorber, respectively. In these expressions E is a kind of "bare" energy difference between the nuclear ground and excited states that would exist if there were no Coulomb effects on the nuclear levels resulting from the surrounding electron distribution and if the nucleus suffered no mass change in the emission and absorption processes. In general, these two energy expressions are not equal because the differences between the environments of the emitting and absorbing nuclei lead to different values for the isomeric and second-order Doppler shifts in the two cases. To achieve an energy balance and hence a resonance condition

experimentally, we impart a velocity to the source relative to the absorber to (first-order) Doppler shift the emitted photons. The change in the energy of the emitted γ ray produced by this step we denote by ΔE_V . The energy balance or resonance condition then becomes an equation for ΔE_V which can be written $\Delta E_V = \Delta E_A$ (isomeric) - ΔE_S (isomeric) + ΔE_A (s.o.D.) - ΔE_S (s.o.D.). The expression for ΔE_V is simply $E (\Delta V/c)$, where ΔV is the velocity imparted to the source and c is the speed of light. In writing this result we can safely neglect the isomeric and second-order Doppler shift corrections to E . The resonance speed imparted to the source is thus given by

$$\begin{aligned} \Delta V_r &= c \left\{ \frac{\Delta E_A(\text{isomeric})}{E} - \frac{\Delta E_S(\text{isomeric})}{E} \right. \\ &\quad \left. + \frac{\Delta E_A(\text{s.o.D.})}{E} - \frac{\Delta E_S(\text{s.o.D.})}{E} \right\} \\ &= \Delta V_A(\text{isomeric}) - \Delta V_S(\text{isomeric}) \\ &\quad + \Delta V_A(\text{s.o.D.}) - \Delta V_S(\text{s.o.D.}) \end{aligned} \quad (4.3.22)$$

where we have expressed the various relative energy shifts as velocity shifts. It is ΔV_r that is measured experimentally, for example, by measuring the intensity of the γ rays transmitted by the absorber as a function of ΔV . The intensity shows a marked decrease when $\Delta V = \Delta V_r$ as given by Eq. (4.3.22).

With the aid of Eq. (4.3.21) we can rewrite Eq. (4.3.22)

$$\begin{aligned} \Delta V_r &= \Delta V_A(\text{isomeric}) - [\Delta V_S(\text{isomeric}) + \Delta V_S(\text{s.o.D.})] \\ &\quad + \frac{3kT}{2M^*c} \left\{ 1 + \frac{1}{20} \frac{M}{M^*} \left(\frac{\Theta}{T} \right)^2 + \dots \right\} \end{aligned} \quad (4.3.23)$$

where

$$\Theta^2 = -\frac{5}{3} \frac{\hbar^2}{k} \frac{1}{3M} \sum_{\alpha} [\bar{\Phi}_{\alpha\alpha}(\iota) + \Delta \bar{\Phi}_{\alpha\alpha}(\iota)] > 0 \quad (4.3.24)$$

If $\Delta \mathcal{G}_{\text{aa}}(\mathcal{L}) = 0$, \mathcal{G} reduces to \mathcal{G}^∞ defined by Eq. (4.2.19).

According to Eq. (4.3.23), if we plot $\Delta V_r - (3kT/2M'c)$ against $1/T$ for the absorber at high temperatures, with the source at constant temperature, we should obtain a straight line, the slope of which determines \mathcal{G} and the intercept of which is

$$\Delta V_A(\text{isomeric}) - \Delta V_S(\text{isomeric}) + \Delta V_S(\text{s.o.D.})$$

If we do this for two different absorbers, using the same source at constant temperature, the difference of the two intercepts gives directly the difference between the isomeric shifts for the two absorber materials independent of the properties of the source and fully corrected for the second-order Doppler shift. Such an experiment has been carried out for Fe^{57} in beryllium and in stainless steel, and the results are shown in Fig. IV.2.²⁶

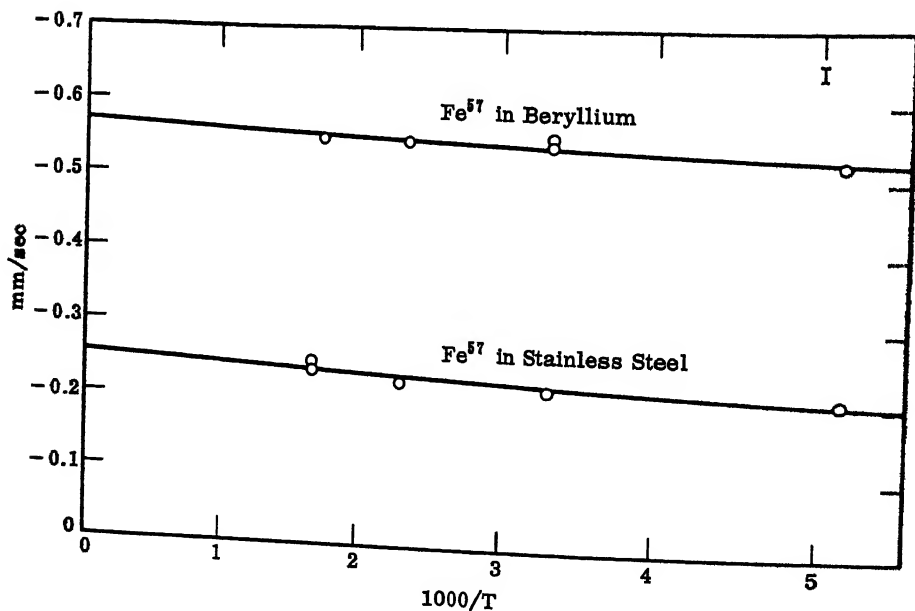


Fig. IV.2 A plot of $\Delta V_r - (3kT/2M'c)$ against the reciprocal temperature for Fe^{57} in beryllium and in stainless steel.

It might be thought that Eq. (4.3.23), which has been obtained for a Bravais crystal, is inapplicable to nonprimitive crystals, such as beryllium. In fact, it is valid for nonprimitive crystals provided we define the characteristic temperature Θ in this case by

$$\Theta^2 = \frac{5}{9} \left(\frac{M'}{M} \right)^2 \left(\frac{\hbar}{k} \right)^2 \left(\frac{d}{d\epsilon} 3rN\Delta\mu_2 \right)$$

where r is the number of atoms in a unit cell. The interpretation of Θ in terms of changes in atomic force constants and masses is somewhat more cumbersome in the case of beryllium than it is for a cubic Bravais lattice. In any event, because of present experimental uncertainties in the determination of Θ , no attempt was made to extract information about the effective force constants between an impurity and its neighbors in either of the two cases studied experimentally.

A least squares fit of the data shown in Fig. IV.2 are given in Table IV.3. The isomeric shift is clearly determined far more accurately

Table IV.3

Absorber	$\Delta V_A(i) - [\Delta V_S(i) - \Delta V_S(\text{s.o.D.})]$	$\Theta(^{\circ}\text{K})$
Fe in beryllium	$-0.570 \pm 0.004, \text{ mm/sec}$	1210 ± 240
Fe in 304 stainless	-0.253 ± 0.003	510 ± 75
Isomeric shift		
Be-stainless	-0.317 ± 0.005	

than the effective characteristic temperature Θ , since the slopes of such curves are much more sensitive to errors in the experimental data than are the intercepts. The results for Θ , however, are not unreasonable. The limiting high-temperature value of the equivalent Debye temperature for beryllium is 900°K .³² Comparison with the result suggests that an iron atom in a beryllium lattice is somewhat more stiffly bound than are the beryllium atoms themselves, but the large uncertainty in Θ makes this suggestion quite tentative. The Debye temperature for stainless steel is not known, but since the values for the constituent elements are

all in the range from 400° to 500° K, the result quoted above is reasonable.

We see finally that nuclear spectroscopy by means of the Mössbauer effect is different from nuclear spectroscopy carried out by more conventional methods in that in the present case it is not possible to measure an energy shift absolutely: only the difference between relative energy shifts for nuclei in different environments can be determined.

V. THERMODYNAMIC FUNCTIONS OF ISOTOPICALLY DISORDERED CRYSTALS

1. Introduction

We have so far been concerned with the effects of isolated impurities on the vibrational properties of crystals. We turn now to study the following problem. Consider a Bravais crystal at every lattice point of which there is either an atom of mass M_1 or an atom of mass M_2 , so that the total number of atoms of mass M_1 is pN , whereas the total number of atoms of mass M_2 is $(1 - p)N$, where N is the total number of atoms in the crystal. We now ask how this mass disorder affects equilibrium properties of our crystal, such as the thermodynamic functions and the frequency spectrum. In the present section we discuss the evaluation of thermodynamic functions and the determination of the frequency spectrum of a disordered crystal is discussed in Sec. VI.

One of the earliest calculations of this type was carried out by Prigogine, Bingen, and Jeener,³³ who studied the effects of disorder on the zero-point energy of isotopic mixtures by perturbation theory. They showed that the interaction between isotopic impurities can lead to the separation of isotopic mixtures into different phases at very low temperatures. Similar conclusions were reached by Chester in a subsequent paper.³⁴

Then Lifshitz and Stepanova³⁵ developed a formal method for calculating the frequency spectrum and Helmholtz free energy of a randomly disordered lattice. The application of their method to obtain numerical results for either of these functions has not yet been carried out.

The evaluation of the change in an arbitrary additive function of the normal mode frequencies of a disordered lattice was studied by Weiss and Maradudin³⁶ who used a perturbation-type expansion which enabled them to discuss the convergence of the resulting series solution.

Recently, the analysis of Weiss and Maradudin has been extended by Chen,³⁷ and it is on this work that the present section is largely based.

In the initial statement of this problem we postulated that our crystal contains exactly pN atoms of mass M_1 and exactly $(1 - p)N$ atoms of mass M_2 . It is, however, convenient to relax this restriction slightly in the following way. We assume instead that the atom at each lattice site has a mass M_1 with probability p , or a mass M_2 with probability $1 - p$. Then the lattice will contain k atoms of mass M_1 with probability

$$\binom{N}{k} p^k (1 - p)^{N-k}$$

If we now consider an ensemble of binomial lattices, the distribution of the number of lattices with k M_1 atoms is found to be approximately Gaussian with mean Np and a standard deviation equal to $[Np(1 - p)]^{1/2}$. As N becomes infinite, the law of large numbers assures us that the probability of a large deviation from the mean number goes to zero. Thus the results we shall obtain apply, to a very good approximation, to the case of a crystal containing exactly pN atoms of mass M_1 randomly distributed over the lattice sites.

Although we have formulated our problem in the preceding paragraphs in terms of a Bravais crystal, this is not a necessary restriction. However, in considering nonprimitive crystals we make one restriction, dictated primarily by a consideration of the applications to which the results of the present analysis will be put. The restriction is that only one of the sublattices is disordered, whereas the remaining sublattices are perfectly ordered. This assumption corresponds to either of the following physical situations: a mixed alkali-halide crystal or a diatomic crystal in which some atoms of one of the constituent species are substituted by one of its isotopes. An example of the former situation would be an NaCl crystal to which we add some KCl substitutionally. The sublattice of chloride ions will remain perfectly ordered, but the sublattice of positive ions is now disordered. An example of the latter situation would be a crystal of LiF in which some of the ${}^7\text{Li}^+$ ions are replaced randomly by ${}^6\text{Li}^+$ ions.

Strictly speaking, the results we shall obtain in this and the next few sections are valid only for isotopically disordered crystals. However, this is probably too severe a restriction, and our results probably hold for alloys or mixtures of chemically similar elements.

We conclude this section by deriving the eigenvalue equation for the normal mode frequencies of an isotopically disordered crystal in a form that is convenient for our subsequent discussion.

The equations of motion of an isotopically disordered crystal can be written

$$M_{\kappa} \ddot{u}_{\kappa} \left(\begin{smallmatrix} \ell \\ \kappa \end{smallmatrix} \right) = - \sum_{\ell' \kappa' \beta} \Phi_{\alpha\beta} \left(\begin{smallmatrix} \ell \ell' \\ \kappa \kappa' \end{smallmatrix} \right) u_{\beta} \left(\begin{smallmatrix} \ell' \\ \kappa' \end{smallmatrix} \right) \quad (5.1.1)$$

We rewrite these equations

$$\ddot{u}_{\alpha} \left(\begin{smallmatrix} \ell \\ \kappa \end{smallmatrix} \right) = - \frac{1}{M_{\kappa}} \sum_{\ell' \kappa' \beta} \Phi_{\alpha\beta} \left(\begin{smallmatrix} \ell \ell' \\ \kappa \kappa' \end{smallmatrix} \right) u_{\beta} \left(\begin{smallmatrix} \ell' \\ \kappa' \end{smallmatrix} \right) - \left(\frac{1}{M_{\kappa}} - \frac{1}{M_{\kappa}} \right) \sum_{\ell' \kappa' \beta} \Phi_{\alpha\beta} \left(\begin{smallmatrix} \ell \ell' \\ \kappa \kappa' \end{smallmatrix} \right) u_{\beta} \left(\begin{smallmatrix} \ell' \\ \kappa' \end{smallmatrix} \right) \quad (5.1.2)$$

In these equations M_{κ} is a mean mass of the atoms of the κ -th sublattice. We have some freedom in choosing M_{κ} . In some cases it is convenient to define M_{κ} as the arithmetic mean of the masses of the atoms on κ -th sublattice,

$$M_{\kappa} = \frac{1}{N} \sum_{\ell} M_{\kappa}^{\ell} \quad (5.1.3)$$

On the other hand, in cases where we are interested in studying the effects of a small concentration of impurities on the dynamic properties of crystals, it is convenient to choose for M_{κ} the mass of the κ -th type atom in the unperturbed crystal. The second term on the right side of Eq. (5.1.2) will be regarded as a perturbation on the remainder of this equation, which is the equation of motion for the mean mass crystal.

We now carry out the normal coordinate transformation to diagonalize the unperturbed (mean mass) equations of motion,

$$u_{\alpha} \left(\begin{smallmatrix} \ell \\ \kappa \end{smallmatrix} \right) = \frac{1}{\sqrt{NM_{\kappa}}} \sum_{\kappa} e_{\alpha} \left(\begin{smallmatrix} \kappa \\ j \end{smallmatrix} \middle| \begin{smallmatrix} k \\ j \end{smallmatrix} \right) Q \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) e^{2\pi i \mathbf{k} \cdot \mathbf{x}(\ell)} \quad (5.1.4)$$

If we substitute Eq. (5.1.4) into Eq. (5.1.2) and use the orthonormality of the eigenvectors $\{e(\kappa) | \mathbf{k}\}$ and of the functions

$$\left\{ \frac{1}{\sqrt{N}} e^{2\pi i \mathbf{k} \cdot \mathbf{x}(\ell)} \right\}$$

we obtain the equation of motion satisfied by $\{Q(\mathbf{k})\}$:

$$\ddot{Q}\left(\begin{smallmatrix} \mathbf{k} \\ j \end{smallmatrix}\right) + \omega_j^2(\mathbf{k}) Q\left(\begin{smallmatrix} \mathbf{k} \\ j \end{smallmatrix}\right) = - \sum_{\mathbf{k}', j'} \Phi\left(\begin{smallmatrix} \mathbf{k} & \mathbf{k}' \\ j & j' \end{smallmatrix}\right) Q\left(\begin{smallmatrix} \mathbf{k}' \\ j' \end{smallmatrix}\right) \quad (5.1.5)$$

where the $\{\omega_j(\mathbf{k})\}$ are now the normal mode frequencies of the mean mass lattice, and

$$\begin{aligned} \Phi\left(\begin{smallmatrix} \mathbf{k} & \mathbf{k}' \\ j & j' \end{smallmatrix}\right) &= \frac{\omega_{j'}^2(\mathbf{k}')}{N} \sum_{\ell} e^{*}\left(\kappa \middle| \begin{smallmatrix} \mathbf{k} \\ j \end{smallmatrix}\right) \cdot e\left(\kappa \middle| \begin{smallmatrix} \mathbf{k}' \\ j' \end{smallmatrix}\right) \\ &\times \left(\frac{M_{\kappa}}{M_{\kappa}'} - 1 \right) e^{-2\pi i (\mathbf{k} - \mathbf{k}') \cdot \mathbf{x}(\ell)} \end{aligned} \quad (5.1.6)$$

If we assume a harmonic time dependence for $Q(\mathbf{k})$,

$$\ddot{Q}\left(\begin{smallmatrix} \mathbf{k} \\ j \end{smallmatrix}\right) = -\omega_j^2 Q\left(\begin{smallmatrix} \mathbf{k} \\ j \end{smallmatrix}\right) \quad (5.1.7)$$

then the equation for the coordinates $\{Q(\mathbf{k})\}$ becomes

$$\sum_{\mathbf{k}', j'} \left\{ \Phi\left(\begin{smallmatrix} \mathbf{k} & \mathbf{k}' \\ j & j' \end{smallmatrix}\right) + [\omega_j^2(\mathbf{k}) - \omega_{j'}^2] \Delta(\mathbf{k} - \mathbf{k}') \delta_{jj'} \right\} Q\left(\begin{smallmatrix} \mathbf{k}' \\ j' \end{smallmatrix}\right) = 0 \quad (5.1.8)$$

The condition that this set of equations have a nontrivial solution is that the determinant of the coefficients vanish, and the resulting equation is the secular equation for the normal mode frequencies of the disordered crystal,

$$\left| \hat{\epsilon} \begin{pmatrix} k & k' \\ j & j' \end{pmatrix} + [\omega_j^2(k) - \omega^2] \Delta(k - k') \delta_{jj'} \right| = 0 \quad (5.1.9)$$

2. The Helmholtz Free Energy of an Isotopically Disordered Crystal

The starting point for our calculation of the Helmholtz free energy of an isotopically disordered crystal is the contour integral formula

$$F(T) = \frac{1}{2\pi i} \int kT \ln \left\{ 2 \sinh \frac{\hbar}{2kT} \right\} d \ln |D(z)| \quad (5.2.1)$$

where $|D(\omega)|$ is the secular determinant whose zeroes are the normal mode frequencies of the disordered crystal. This determinant is given explicitly by Eq. (5.1.9) which we rewrite

$$\left| D \begin{pmatrix} k & k' \\ j & j' \end{pmatrix}; \omega^2 \right| = 0 \quad (5.2.2)$$

where $D(kj; k'j'; \omega^2)$ is the (kj) $(k'j')$ element of a matrix $D(\omega^2)$ and is given by

$$D \begin{pmatrix} k & k' \\ j & j' \end{pmatrix}; \omega^2 = \hat{\epsilon} \begin{pmatrix} k & k' \\ j & j' \end{pmatrix} + [\omega_j^2(k) - \omega^2] \Delta(k - k') \delta_{jj'} \quad (5.2.3)$$

In writing Eqs. (5.2.2) and (5.2.3) we have expressed explicitly what has been stated only implicitly up to now, namely that the secular determinant of a crystal depends on ω only through the function ω^2

Let us write $|D(\omega^2)|$ formally as

$$|D(\omega^2)| = \prod_{\lambda} (\omega_{\lambda}^2 - \omega^2) \quad (5.2.4)$$

where $\{\omega_{\lambda}\}$ are the normal mode frequencies of the disordered crystal, and λ runs over all $3rN$ degrees of freedom in the crystal. From Eq. (5.2.4) we obtain

$$\frac{d}{d\omega} \ln |D(\omega^2)| = 2\omega \sum_{\lambda} \frac{1}{\omega_{\lambda}^2 - \omega^2} \quad (5.2.5)$$

and this is seen to be the same as

$$\begin{aligned} \frac{d}{d\omega} \ln |D(\omega^2)| &= 2\omega \operatorname{Tr} D^{-1}(\omega^2) \\ &= -2\omega \sum_{kj} D^{-1} \left(\begin{matrix} k & k \\ j & j \end{matrix} ; \omega^2 \right) \end{aligned} \quad (5.2.6)$$

where $D^{-1}(\omega^2)$ is the matrix which is inverse to $D(\omega^2)$

The elements of $D^{-1}(\omega^2)$ satisfy the equation

$$\sum_{k'j'} D \left(\begin{matrix} k & k' \\ j & j' \end{matrix} ; \omega^2 \right) D^{-1} \left(\begin{matrix} k & k' \\ j & j' \end{matrix} ; \omega^2 \right) = \Delta(k - k') \delta_{jj'} \quad (5.2.7)$$

If we substitute Eq. (5.2.3) into this equation, we find that

$$D^{-1} \left(\begin{matrix} k & k' \\ j & j' \end{matrix} ; \omega^2 \right) = \frac{\Delta(k - k') \delta_{jj'}}{\omega_j^2(k) - \omega^2} - \frac{1}{\omega_j^2(k) - \omega^2}$$

$$\sum_{k_1 j_1} \Phi \left(\begin{matrix} k & k_1 \\ j & j_1 \end{matrix} \right) D^{-1} \left(\begin{matrix} k_1 & k' \\ j_1 & j' \end{matrix} ; \omega^2 \right) \quad (5.2.8)$$

We now proceed to solve this equation by iteration. We obtain the formal expansion

$$\begin{aligned} D^{-1} \left(\begin{matrix} k & k' \\ j & j' \end{matrix} ; \omega^2 \right) &= \frac{\Delta(k - k') \delta_{jj'}}{d_j(k)} - \frac{1}{d_j(k)} \\ &\times \Phi \left(\begin{matrix} k & k' \\ j & j' \end{matrix} \right) \frac{1}{d_{j_1}(k')} + \frac{1}{d_j(k)} \\ &\times \sum_{k_1 j_1} \Phi \left(\begin{matrix} k & k_1 \\ j & j_1 \end{matrix} \right) \frac{1}{d_{j_1}(k_1')} \Phi \left(\begin{matrix} k_1 & k' \\ j_1 & j' \end{matrix} \right) \frac{1}{d_{j_1}(k')} - \dots \end{aligned} \quad (5.2.9)$$

where we have introduced

$$d_j(k) = \omega_j^2(k) - \omega^2 \quad (5.2.10)$$

Up to this point our results are valid for any configuration of the isotopic impurities over the lattice sites. However, we have no a priori knowledge of what the configuration is, and the best we can do is to average $D^{-1}(k j; k' j'; \omega^2)$ over all possible configurations of the two kinds of atoms over the lattice sites of the crystal that are compatible with the occupation probabilities stated in the preceding section, assigning the same weight to each configuration.

We thus consider instead of Eq. (5.2.9) the expansion

$$\begin{aligned} \left\langle D^{-1} \left(\begin{smallmatrix} k & k' \\ j & j' \end{smallmatrix}; \omega^2 \right) \right\rangle_A &= \frac{\Delta(k - k') \delta_{jj'}}{d_j(k)} - \frac{1}{d_j(k)} \left\{ \left\langle \tilde{\Phi} \left(\begin{smallmatrix} k & k' \\ j & j' \end{smallmatrix} \right) \right\rangle_A \right. \\ &\quad \left. - \sum_{k_1 j_1} \left\langle \tilde{\Phi} \left(\begin{smallmatrix} k & k_1 \\ j & j_1 \end{smallmatrix} \right) \frac{1}{d_{j_1}(k_1)} \tilde{\Phi} \left(\begin{smallmatrix} k_1 & k' \\ j_1 & j' \end{smallmatrix} \right) \right\rangle_A + \dots \right\} \frac{1}{d_{j_1}(k')} \end{aligned}$$

(5.2.11)

We now have to consider the nature of the averages of products of the $\tilde{\Phi}$ coefficients. It is at this point that we invoke the assumption we have discussed in the preceding section, namely that in dealing with nonprimitive crystals we assume that only one of the sublattices is disordered. Let us denote the sublattice which is disordered by +. The coefficient $\tilde{\Phi}(k_j; k'_{j'})$ now assumes the form

$$\begin{aligned} \tilde{\Phi} \left(\begin{smallmatrix} k & k' \\ j & j' \end{smallmatrix} \right) &= \frac{\omega_{j'}^2(k')}{N} \sum_{\ell} \left(\frac{M_+}{M_+^{\ell}} - 1 \right) e^{*} \left(+ \mid \begin{smallmatrix} k \\ j \end{smallmatrix} \right) \cdot e \left(+ \mid \begin{smallmatrix} k' \\ j' \end{smallmatrix} \right) \\ &\times e^{-2\pi i(k - k') \cdot x(\ell)} \end{aligned}$$

(5.2.12)

where M_+ is a mean mass for the + sublattice. We now introduce the notation

$$\left\langle \left(\frac{M_+}{M_+^{\ell}} - 1 \right)^n \right\rangle_A = \sigma_n$$

(5.2.13)

We can now write out the averages of products of several $\tilde{\Phi}$'s. We have first of all

$$\left\langle \hat{\Phi} \begin{pmatrix} k & k' \\ j & j' \end{pmatrix} \right\rangle_A = \sigma_1 \frac{\omega_{j'}^2(k')}{N} e^* \left(+ \mid \begin{matrix} k \\ j \end{matrix} \right) \cdot e \left(+ \mid \begin{matrix} k' \\ j' \end{matrix} \right) \\ \times \sum_{\ell} e^{-2\pi i(k-k') \cdot x(\ell)}$$

(5. 2. 14)

since the operations of taking the average and summing over ℓ commute. Consequently, we obtain

$$\left\langle \hat{\Phi} \begin{pmatrix} k & k' \\ j & j' \end{pmatrix} \right\rangle_A = \sigma_1 \omega_{j'}^2(k') e^* \left(+ \mid \begin{matrix} k \\ j \end{matrix} \right) \cdot e \left(+ \mid \begin{matrix} k' \\ j' \end{matrix} \right) \Delta(k - k')$$

(5. 2. 15)

where we have used the relation

$$\sum_{\ell} e^{2\pi i k \cdot x(\ell)} = N \Delta(k)$$

(5. 2. 16)

The average of a product of two $\hat{\Phi}$'s is formally

$$\left\langle \hat{\Phi} \begin{pmatrix} k & k_1 \\ j & j_1 \end{pmatrix} \hat{\Phi} \begin{pmatrix} k_1 & k' \\ j_1 & j'_1 \end{pmatrix} \right\rangle_A = \frac{\omega_{j_1}^2(k_1) \omega_{j'_1}^2(k')}{N^2} e^* \left(+ \mid \begin{matrix} k \\ j \end{matrix} \right) \cdot e \left(+ \mid \begin{matrix} k_1 \\ j_1 \end{matrix} \right) \\ \times e^* \left(+ \mid \begin{matrix} k_1 \\ j_1 \end{matrix} \right) \cdot e \left(+ \mid \begin{matrix} k' \\ j'_1 \end{matrix} \right) \sum_{\ell_1 \ell_2} \left\langle \left(\frac{M_+}{M_+^{\ell_1}} - 1 \right) \left(\frac{M_+}{M_+^{\ell_2}} - 1 \right) \right\rangle_A$$

$$\times e^{-2\pi i(k-k_1) \cdot x(\ell_1) - 2\pi i(k_1-k') \cdot x(\ell_2)}$$

(5. 2. 17)

We now break up the sums over ℓ_1 and ℓ_2 into those terms for which $\ell_1 = \ell_2$ and those terms for which $\ell_1 \neq \ell_2$. The average of the double sum thus becomes

$$\begin{aligned} & \sigma_2 \sum_{\ell_1} e^{-2\pi i(\mathbf{k}-\mathbf{k}') \cdot \mathbf{x}(\ell)} \\ & + \sigma_1^2 \sum_{\ell_1 \neq \ell_2} e^{-2\pi i(\mathbf{k}-\mathbf{k}_1) \cdot \mathbf{x}(\ell_1)} e^{-2\pi i(\mathbf{k}_1-\mathbf{k}') \cdot \mathbf{x}(\ell_2)} \end{aligned} \quad (5.2.18)$$

since in the second term there is no correlation between the masses at the distinct lattice points $\mathbf{x}(\ell_1)$ and $\mathbf{x}(\ell_2)$. To evaluate the second sum we add in and subtract off the terms with $\ell_1 = \ell_2$. The expression in (5.2.18) becomes

$$\sigma_2 N \Delta(\mathbf{k}-\mathbf{k}') + \sigma_1^2 [N^2 \Delta(\mathbf{k}-\mathbf{k}_1) \Delta(\mathbf{k}_1-\mathbf{k}') - N \Delta(\mathbf{k}-\mathbf{k}')] \quad (5.2.19)$$

and we finally obtain

$$\begin{aligned} & \left\langle \hat{\Phi} \begin{pmatrix} \mathbf{k} & \mathbf{k}_1 \\ j & j_1 \end{pmatrix} \hat{\Phi} \begin{pmatrix} \mathbf{k}_1 & \mathbf{k}' \\ j_1 & j' \end{pmatrix} \right\rangle_A = (\sigma_2 - \sigma_1^2) \frac{\omega_{j_1}^2(\mathbf{k}_1) \omega_{j'}^2(\mathbf{k}')}{N} \\ & \times e^{*} \left(+ \left| \begin{matrix} \mathbf{k} \\ j \end{matrix} \right. \right) \cdot e \left(+ \left| \begin{matrix} \mathbf{k}_1 \\ j_1 \end{matrix} \right. \right) e^{*} \left(+ \left| \begin{matrix} \mathbf{k}_1 \\ j_1 \end{matrix} \right. \right) \cdot e \left(+ \left| \begin{matrix} \mathbf{k}' \\ j' \end{matrix} \right. \right) \Delta(\mathbf{k}-\mathbf{k}') \\ & + \sigma_1^2 \omega_{j_1}^2(\mathbf{k}_1) \omega_{j'}^2(\mathbf{k}') e^{*} \left(+ \left| \begin{matrix} \mathbf{k} \\ j \end{matrix} \right. \right) \cdot e \left(+ \left| \begin{matrix} \mathbf{k}_1 \\ j_1 \end{matrix} \right. \right) \\ & \times e^{*} \left(+ \left| \begin{matrix} \mathbf{k}_1 \\ j_1 \end{matrix} \right. \right) \cdot e \left(+ \left| \begin{matrix} \mathbf{k}' \\ j' \end{matrix} \right. \right) \Delta(\mathbf{k}-\mathbf{k}') \Delta(\mathbf{k}-\mathbf{k}_1) \end{aligned} \quad (5.2.20)$$

Averages of products of more \bar{f} s are obtained in the same way. We split the multiple sums over i_1, i_2, \dots, i_n into a sum with all indices unequal to each other, plus sums with two indices equal in the various possible combinations, plus sums with three equal, plus sums with two pairs set equal separately, and so forth, down to a single sum with all indices equal. In each sum indices that are not explicitly set equal are restricted to be different from each other. We may now average each sum over the quantities associated with each index separately. After the averages are carried out and all the dependences on the masses have been removed outside the sums, we then add in and subtract off terms in the summand in such a way that the sums over the distinct indices can be carried out independently of each other over all N values that each index can assume.

The results expressed by Eqs. (5.2.15) and (5.2.20) already show the essential features of all higher-order averages. We see first of all that each average is diagonal in the wave vectors k and k' . This is a general result which follows from the translational invariance of the crystal after we have performed the configuration average: every lattice site on the disordered sublattice is equivalent to every other site on the same sublattice. The fact that j and j' are not necessarily equal is referred to as polarization mixing. In addition we see that averaging over all configurations has the effect of forcing some of the summation variables $\{k_i\}$ appearing in the expansion (5.2.11) to equal k .

We can give a picturesque description of the various terms which appear in the expansion (5.2.11). A free phonon of momentum k and polarization j (if we may use this term in a purely classical problem) described by the function $d_j^{-1}(k)$ is scattered by an impurity at some lattice site, propagates freely again for a while, is scattered by the same or by a different impurity, propagates freely again, and so on until it emerges with the same "momentum" k , but with a different polarization. Each Δ -function in the averages (5.2.15) and (5.2.20) corresponds to the scattering of our phonon from a different lattice point.

We can depict graphically the various terms which appear in Eq. (5.2.11). Each free phonon propagator $d_j^{-1}(k)$ is depicted by a solid line directed upward, whereas each interaction with an impurity is denoted by a dashed line starting at a point that corresponds to the lattice site at which the interaction occurs and joins the free phonon line in the order in which it occurs in the expansion (5.2.11). All interaction lines corresponding to the scattering of a phonon by the same impurity emanate from the same point. These are essentially the diagrams introduced by Langer³⁸ in his analysis of the disordered chain. In Fig. V.1 is shown the diagram corresponding to the first-order term in Eq. (5.2.11) and the average given by Eq. (5.2.15). In Fig. V.2 are shown the diagrams which

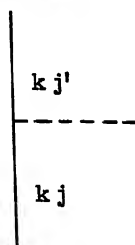


Fig. V.1 Diagram corresponding to the first-order term in Eq. (5.2.11).

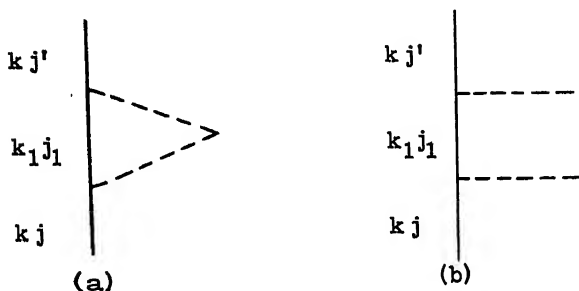


Fig. V.2 Diagrams corresponding to the second-order terms in Eq. (5.2.11).

correspond to the two second-order terms in Eq. (5.2.11). Figure V.2a corresponds to the first term of the average in Eq. (5.2.20) and describes the double scattering of a phonon by the same impurity; Fig. V.2b corresponds to the second term of the average in Eq. (5.2.20) and describes the single scattering of a phonon by two different impurities. The third-order diagrams are shown in Fig. V.3.

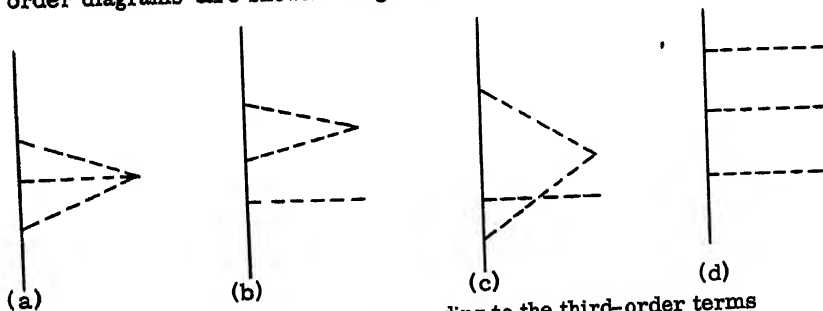


Fig. V.3 Diagrams corresponding to the third-order terms in Eq. (5.2.11).

A typical diagram contributing to $\langle D^{-1}(k j; k' j'; \omega^2) \rangle_A$ has the following structure. A free phonon line ($k j$) enters the diagram at the bottom, interactions of all degrees of complexity occur, and a free phonon line ($k j'$) leaves the diagram at the top. From Figs. V. 1, V. 2, and V. 3, we see that these diagrams fall into two categories: those that can be separated into two unconnected parts by cutting a single solid line; and those that cannot. The former type of diagram is called "improper"; the latter type is called "proper." Examples of proper diagrams are Figs. V. 1a, V. 2a, V. 3a, and V. 3c, whereas Figs. V. 2b, V. 3b, and V. 3d depict improper diagrams.

The average matrix element $\langle D^{-1}(k j; k' j'; \omega^2) \rangle_A$ can be expressed in terms of contributions from proper diagrams only. This is convenient because it greatly reduces the number of diagrams whose contributions have to be considered. Let us introduce a function $G(k j j')$ that is defined as the sum of the contributions to the expansion in curly brackets in Eq. (5.2.11) associated with proper diagrams only. $G(k j j')$ is called the proper self-energy. In terms of this function we obtain

$$\begin{aligned} \langle D^{-1} \left(\begin{smallmatrix} k & k' \\ j & j' \end{smallmatrix}; \omega^2 \right) \rangle_A &= \Delta(k - k') \left\{ \frac{\delta_{jj'}}{d_j(k)} - \frac{1}{d_j(k)} \right. \\ &\times \left[G \left(\begin{smallmatrix} k \\ jj' \end{smallmatrix} \right) - \sum_{j_1} G \left(\begin{smallmatrix} k \\ jj_1 \end{smallmatrix} \right) \frac{1}{d_{j_1}(k)} G \left(\begin{smallmatrix} k \\ j_1 j' \end{smallmatrix} \right) \right. \\ &+ \sum_{j_1 j_2} G \left(\begin{smallmatrix} k \\ jj_1 \end{smallmatrix} \right) \frac{1}{d_{j_1}(k)} G \left(\begin{smallmatrix} k \\ j_1 j_2 \end{smallmatrix} \right) \frac{1}{d_{j_2}(k)} G \left(\begin{smallmatrix} k \\ j_2 j' \end{smallmatrix} \right) - \dots \left. \right] \\ &\times \left. \frac{1}{d_{j'}(k)} \right\} = \Delta(k - k') D^{-1} \left(\begin{smallmatrix} k \\ jj' \end{smallmatrix}; \omega^2 \right) \end{aligned}$$

(5.2.21)

$D^{-1}(k j j'; \omega^2)$ is thus given by the solution of the $3r \times 3r$ matrix equation

$$\begin{aligned}
D^{-1} \left(\begin{matrix} k \\ jj' \end{matrix} ; \omega^2 \right) &= \frac{\delta_{jj'}}{d_j(k)} \\
&- \frac{1}{d_j(k)} \sum_{j_1} G \left(\begin{matrix} k \\ jj_1 \end{matrix} \right) D^{-1} \left(\begin{matrix} k \\ j_1 j' \end{matrix} ; \omega^2 \right)
\end{aligned}
\tag{5.2.22}$$

If it were not for polarization mixing we could solve this set of equations directly.

In a sense, we have now formally solved the problem of determining the Helmholtz free energy of an isotopically disordered crystal. From Eq. (5.2.22) and the definition of the proper self-energy $G(k, jj')$ we can solve for $D^{-1}(k, jj'; \omega^2)$ to some preassigned degree of accuracy. This function, together with Eqs. (5.2.6) and (5.2.1), determines the free energy. However, as it stands, a great deal of computational effort would undoubtedly be required to translate our formal expressions into numerical results. It is, therefore, of considerable interest to see how much farther we can proceed in a purely analytic manner before we are finally forced into numerical calculations.

To proceed beyond this point we need to know something about the function $G(k, jj')$. In the case of a nonprimitive crystal, we shall only be interested, in subsequent applications, in a result that is correct to terms linear in the concentration of impurities. For this purpose it is convenient to define the mean mass M_+ as the mass of the ions comprising the + sublattice in the unperturbed crystal. If the mass of the substitutional impurity is denoted by M_+^i , then we have

$$\begin{aligned}
M_+^i &= M_+ \quad \text{with probability } 1 - p \\
&= M_+^i \quad \text{with probability } p
\end{aligned}
\tag{5.2.23}$$

The average σ_n defined by Eq. (5.2.13) is given in this case by

$$\sigma_n = \left\langle \left(\frac{M_+}{M_+^i} - 1 \right)^n \right\rangle_A = p \left(\frac{M_+}{M_+^i} - 1 \right)^n = p \lambda^n
\tag{5.2.24}$$

We now determine $G(k|j|j')$ correct to terms linear in p . To terms linear in p we find

$$\left\langle \sharp \begin{pmatrix} k & k \\ j & j' \end{pmatrix} \right\rangle_A = \sigma_1 w_{j'}^2(k) e^* \left(+ \mid \begin{pmatrix} k \\ j \end{pmatrix} \right) \cdot e \left(+ \mid \begin{pmatrix} k \\ j' \end{pmatrix} \right) \quad (5.2.25a)$$

$$\begin{aligned} \left\langle \sharp \begin{pmatrix} k & k_1 \\ j & j_1 \end{pmatrix} \sharp \begin{pmatrix} k_1 & k \\ j_1 & j' \end{pmatrix} \right\rangle_A &= \sigma_2 \frac{w_{j_1}^2(k) w_{j'}^2(k)}{N} \\ &\times \left[e^* \left(+ \mid \begin{pmatrix} k \\ j \end{pmatrix} \right) \cdot e \left(+ \mid \begin{pmatrix} k_1 \\ j_1 \end{pmatrix} \right) \right] \left[e^* \left(+ \mid \begin{pmatrix} k_1 \\ j_1 \end{pmatrix} \right) \cdot e \left(+ \mid \begin{pmatrix} k \\ j' \end{pmatrix} \right) \right] \end{aligned} \quad (5.2.25b)$$

In general, the term linear in p in the average

$$\left\langle \sharp \begin{pmatrix} k & k_1 \\ j & j_1 \end{pmatrix} \dots \sharp \begin{pmatrix} k_{n-1} & k \\ j_{n-1} & j' \end{pmatrix} \right\rangle_A$$

arises from those terms in which the n summation variables are set equal to each other

$$x(l_1) = x(l_2) = \dots = x(l_n) = x(l)$$

That is, it is the term proportional to σ_n in the average of a product of $n \sharp$'s. Physically this result corresponds to the repeated scattering of a phonon by the same impurity atom.

The function $G(k|j|j')$ is given explicitly by

$$\begin{aligned} G \left(\begin{pmatrix} k \\ j|j' \end{pmatrix} \right) &= \left\langle \sharp \begin{pmatrix} k & k \\ j & j' \end{pmatrix} \right\rangle_A \\ &- \sum_{k_1 j_1} \frac{\left\langle \sharp \begin{pmatrix} k & k_1 \\ j & j_1 \end{pmatrix} \sharp \begin{pmatrix} k_1 & k \\ j_1 & j' \end{pmatrix} \right\rangle_A}{d_{j_1}(k_1)} + \dots \end{aligned} \quad (5.2.26)$$

where the prime restricts the sum to contributions from proper diagrams only. Substituting Eqs. (5.2.25) into Eq. (5.2.26) we find

$$G \left(\begin{matrix} k \\ j \ j' \end{matrix} \right) = \omega_{j'}^2(k) p\lambda \left\{ \sum_{\alpha} e_{\alpha}^* \left(+ \mid \begin{matrix} k \\ j \end{matrix} \right) e_{\alpha} \left(+ \mid \begin{matrix} k \\ j' \end{matrix} \right) \right. \\ \left. - \frac{\lambda}{N} \sum_{\alpha\beta} \sum_{k_1 j_1} \frac{e_{\alpha}^* \left(+ \mid \begin{matrix} k \\ j \end{matrix} \right) e_{\alpha} \left(+ \mid \begin{matrix} k_1 \\ j_1 \end{matrix} \right) \omega_{j_1}^2(k_1) e_{\beta}^* \left(+ \mid \begin{matrix} k_1 \\ j_1 \end{matrix} \right) e_{\beta} \left(+ \mid \begin{matrix} k \\ j' \end{matrix} \right)}{d_{j_1}(k_1)} + \dots \right\} \quad (5.2.27)$$

The diagrammatic representation of this series is given in Fig. V.4.

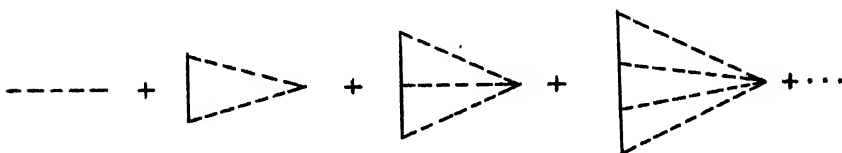


Fig. V.4 Diagrammatic representation for the expansion which yields the proper self-energy function $G(k \ j \ j')$ correct to terms linear in p .

Let us introduce the $3r \times 3r$ matrix K whose elements are given by

$$K_{\alpha\beta} = - \frac{\lambda}{N} \sum_{kj} \frac{e_{\alpha} \left(+ \mid \begin{matrix} k \\ j \end{matrix} \right) \omega_j^2(k) e_{\beta}^* \left(+ \mid \begin{matrix} k \\ j \end{matrix} \right)}{d_j(k)} \quad (5.2.28)$$

In terms of this matrix we find

$$G \left(\begin{matrix} k \\ j \ j' \end{matrix} \right) = \omega_{j'}^2(k) p\lambda \sum_{\alpha\beta} e_{\alpha}^* \left(+ \mid \begin{matrix} k \\ j \end{matrix} \right) \left\{ \delta_{\alpha\beta} + K_{\alpha\beta} \right.$$

$$\begin{aligned}
 & + \sum_Y K_{\alpha Y} K_{Y\beta} + \dots \left\} e_{\beta} \left(+ \left| \begin{smallmatrix} k \\ j' \end{smallmatrix} \right. \right) \\
 & = p\lambda \omega_{j'}^2(k) \sum_{\alpha\beta} e_{\alpha}^* \left(+ \left| \begin{smallmatrix} k \\ j \end{smallmatrix} \right. \right) [I - K]_{\alpha\beta}^{-1} e_{\beta} \left(+ \left| \begin{smallmatrix} k \\ j' \end{smallmatrix} \right. \right)
 \end{aligned}$$

(5. 2. 29)

We can rewrite $K_{\alpha\beta}$ as

$$K_{\alpha\beta} = -\lambda \delta_{\alpha\beta} + \frac{\lambda \omega^2}{N} \sum_{kj} \frac{e_{\alpha} \left(+ \left| \begin{smallmatrix} k \\ j \end{smallmatrix} \right. \right) e_{\beta}^* \left(+ \left| \begin{smallmatrix} k \\ j \end{smallmatrix} \right. \right)}{\omega^2 - \omega_j^2(k)}$$

(5. 2. 30)

By using the arguments of Sec. IV-1, we can assert that $K_{\alpha\beta}$ is isotropic for cubic crystals in which every atom is at a center of inversion symmetry

$$\begin{aligned}
 K_{\alpha\beta} &= \lambda \delta_{\alpha\beta} \left\{ -1 + \frac{\omega^2}{3N} \sum_{kj} \frac{e \left(+ \left| \begin{smallmatrix} k \\ j \end{smallmatrix} \right. \right) \cdot e^* \left(+ \left| \begin{smallmatrix} k \\ j \end{smallmatrix} \right. \right)}{\omega^2 - \omega_j^2(k)} \right\} \\
 &= \lambda \delta_{\alpha\beta} K
 \end{aligned}$$

(5. 2. 31)

In this case we obtain finally

$$G \left(\begin{smallmatrix} k \\ jj' \end{smallmatrix} \right) = p\lambda \omega_{j'}^2(k) \frac{e^* \left(+ \left| \begin{smallmatrix} k \\ j \end{smallmatrix} \right. \right) \cdot e \left(+ \left| \begin{smallmatrix} k \\ j' \end{smallmatrix} \right. \right)}{1 - \lambda K}$$

(5. 2. 32)

For Bravais crystals the leading contribution to $G(kjj')$ is given by

$$G^{(1)}\left(\begin{matrix} k \\ jj' \end{matrix}\right) = \sigma_1 w_j^2(k) \Delta(k - k') \delta_{jj'} \quad (5.2.33)$$

where σ_n is now

$$\sigma_n = \left\langle \left(\frac{M}{M_L} - 1 \right)^n \right\rangle_A \quad (5.2.34)$$

and we have used the orthogonality of the eigenvectors corresponding to the same k vector but belonging to different branches. The fact that $G^{(1)}(kjj')$ is diagonal in j and j' suggests the following way of solving Eq. (5.2.22). We divide $G(kjj')$ into its diagonal and nondiagonal parts,

$$G\left(\begin{matrix} k \\ jj' \end{matrix}\right) = G\left(\begin{matrix} k \\ j \end{matrix}\right) \delta_{jj'} + G^N\left(\begin{matrix} k \\ jj' \end{matrix}\right) \quad (5.2.35)$$

We now rewrite Eq. (5.2.22) as

$$\begin{aligned} & \sum_{j_1} \left\{ \delta_{jj_1} + \frac{1}{d_j(k)} G\left(\begin{matrix} k \\ j \end{matrix}\right) \delta_{jj_1} \right\} D^{-1}\left(\begin{matrix} k \\ j_1 j_1' \end{matrix}; \omega^2\right) \\ &= \frac{\delta_{jj'}}{d_j(k)} - \frac{1}{d_j(k)} \sum_{j_1} G^N\left(\begin{matrix} k \\ j j_1 \end{matrix}\right) D^{-1}\left(\begin{matrix} k \\ j_1 j_1' \end{matrix}; \omega^2\right) \end{aligned} \quad (5.2.36)$$

so that

$$D^{-1}\left(\begin{matrix} k \\ jj' \end{matrix}; \omega^2\right) = \frac{\delta_{jj'}}{d_j(k) + G\left(\begin{matrix} k \\ j \end{matrix}\right)} - \frac{1}{d_j(k) + G\left(\begin{matrix} k \\ j \end{matrix}\right)}.$$

$$\sum_{j_1} G^N \left(\begin{matrix} k \\ j j_1 \end{matrix} \right) D^{-1} \left(\begin{matrix} k \\ j_1 j_1' \end{matrix} ; \omega^2 \right) \quad (5.2.37)$$

This equation can now be solved by iteration and we find

$$D^{-1} \left(\begin{matrix} k \\ j j_1' \end{matrix} ; \omega^2 \right) = \frac{\delta_{j j_1'}}{d_{j_1}(k) + G \left(\begin{matrix} k \\ j \end{matrix} \right)} - \frac{1}{d_{j_1}(k) + G \left(\begin{matrix} k \\ j \end{matrix} \right)} G^N \left(\begin{matrix} k \\ j j_1' \end{matrix} \right) \frac{1}{d_{j_1}(k) + G \left(\begin{matrix} k \\ j_1' \end{matrix} \right)} + \dots \quad (5.2.38)$$

Since $G^N \left(\begin{matrix} k \\ j j_1' \end{matrix} \right)$ has no diagonal elements and thus is of second order in the ϵ 's, we find that the diagonal element $D^{-1} \left(\begin{matrix} k \\ j j; \omega^2 \end{matrix} \right)$ is given with an error of fourth order in the ϵ 's by

$$D^{-1} \left(\begin{matrix} k \\ j j; \omega^2 \end{matrix} \right) \cong \frac{\delta_{j j}}{d_j(k) + G \left(\begin{matrix} k \\ j \end{matrix} \right)} = \frac{\delta_{j j}}{\omega_j^2(k) - \omega^2 + G \left(\begin{matrix} k \\ j \end{matrix} \right)} \quad (5.2.39)$$

We thus have the result

$$\frac{d}{d\omega} \ln |D(\omega)| = 2\omega \sum_{k,j} \frac{1}{\omega^2 - \omega_j^2(k) - G \left(\begin{matrix} k \\ j \end{matrix} \right)}$$

$$(5.2.40)$$

For a Bravais crystal $G(k, j)$ is given (to second order in the perturbation) by

$$G\left(\begin{matrix} k \\ j \end{matrix}\right) = \sigma_1 \omega_j^2(k) + \frac{\sigma_2 - \sigma_1^2}{N} \sum_{k_1 j_1} \frac{\omega_j^2(k) \omega_{j_1}^2(k_1) e\left(\begin{matrix} k \\ j \end{matrix}\right) \cdot e\left(\begin{matrix} k_1 \\ j_1 \end{matrix}\right) e\left(\begin{matrix} k_1 \\ j_1 \end{matrix}\right) \cdot e\left(\begin{matrix} k \\ j \end{matrix}\right)}{\omega^2 - \omega_{j_1}^2(k_1)} \quad (5.2.41)$$

For cubic crystals this result reduces to

$$G\left(\begin{matrix} k \\ j \end{matrix}\right) = \sigma_1 \omega_j^2(k) + \frac{(\sigma_2 - \sigma_1^2) \omega_j^2(k)}{3N} \sum_{k_1 j_1} \frac{\omega_{j_1}(k_1)}{\omega^2 - \omega_{j_1}^2(k_1)} \quad (5.2.42)$$

If we assume that

$$\begin{aligned} M^L &= M_1 \text{ with probability } p \\ &= M_2 \text{ with probability } 1 - p \end{aligned} \quad (5.2.43)$$

and that

$$M = pM_1 + (1 - p) M_2 \quad (5.2.44)$$

we find

$$\sigma_1 = p(1 - p) \frac{(M_1 - M_2)^2}{M_1 M_2} \quad (5.2.45a)$$

$$\sigma_2 - \sigma_1^2 = p(1-p) M^2 \left(\frac{M_1 - M_2}{M_1 M_2} \right)^2$$

(5.2.45b)

With this result we obtain

$$\begin{aligned} G \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) &= p(1-p) \frac{(M_1 - M_2)^2}{M_1 M_2} \omega_j^2(k) \\ &+ p(1-p) M^2 \left(\frac{M_1 - M_2}{M_1 M_2} \right)^2 \\ &\times \frac{\omega_j^2(k)}{3N} \sum_{k_1 j_1} \frac{\omega_{j_1}^2(k_1)}{\omega^2 - \omega_{j_1}^2(k_1)} \end{aligned}$$

(5.2.46)

More interesting, however, is the result for $G \left(\begin{smallmatrix} k \\ j j' \end{smallmatrix} \right)$ for Bravais crystals in the limit that the concentration of impurities is so low that we can be satisfied with knowing $G \left(\begin{smallmatrix} k \\ j j' \end{smallmatrix} \right)$ correct to terms linear in the concentration. This result is readily obtained from the analysis leading to Eq. (5.2.29). In fact we can obtain the desired result from Eq. (5.2.29) itself, if we suppress the index + and recall that the eigenvectors are real in this case. In this way we obtain

$$G \left(\begin{smallmatrix} k \\ j j' \end{smallmatrix} \right) = p \lambda \omega_j^2(k) \sum_{\alpha \beta} e_{\alpha} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) [I - K]_{\alpha \beta}^{-1} e_{\beta} \left(\begin{smallmatrix} k \\ j' \end{smallmatrix} \right)$$

(5.2.47)

with

$$K_{\alpha \beta} = -\lambda \delta_{\alpha \beta} + \frac{\lambda \omega^2}{N} \sum_{k j} \frac{e_{\alpha} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) e_{\beta} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right)}{\omega^2 - \omega_j^2(k)}$$

(5.2.48)

where now

$$\lambda = \frac{M}{M'} - 1 = \frac{\epsilon}{1 - \epsilon} \quad (5.2.49)$$

In the special case of a cubic Bravais crystal these results simplify to

$$\begin{aligned} K_{\alpha\beta} &= \lambda \delta_{\alpha\beta} \left[-1 + \frac{\omega^2}{3N} \sum_{\mathbf{k} \mathbf{j}} \frac{1}{\omega^2 - \omega_{\mathbf{j}}^2(\mathbf{k})} \right] \\ &= \lambda \delta_{\alpha\beta} g \end{aligned} \quad (5.2.50)$$

$$G \begin{pmatrix} \mathbf{k} \\ \mathbf{j} \mathbf{j}' \end{pmatrix} = p \lambda \omega_{\mathbf{j}}^2(\mathbf{k}) \frac{\delta_{\mathbf{j} \mathbf{j}'}}{1 - \lambda g} = \delta_{\mathbf{j} \mathbf{j}'} \mathcal{G} \begin{pmatrix} \mathbf{k} \\ \mathbf{j} \end{pmatrix} \quad (5.2.51)$$

where we have used the orthonormality of eigenvectors belonging to the same \mathbf{k} but to different \mathbf{j} values.

Since $G(\mathbf{k} \mathbf{j} \mathbf{j}')$ is diagonal in \mathbf{j} and \mathbf{j}' to terms linear in p , we can solve Eq. (5.2.22) directly with the result

$$D^{-1} \begin{pmatrix} \mathbf{k} \\ \mathbf{j} \mathbf{j}' ; \omega^2 \end{pmatrix} = \delta_{\mathbf{j} \mathbf{j}'} \frac{1}{\omega_{\mathbf{j}}^2(\mathbf{k}) - \omega_{\mathbf{j}}^2 + \mathcal{G} \begin{pmatrix} \mathbf{k} \\ \mathbf{j} \end{pmatrix}} \quad (5.2.52)$$

We are finally led to the result that, for cubic Bravais crystals,

$$\frac{d}{d\omega} \ln |D(\omega^2)| = 2\omega \sum_{\mathbf{k} \mathbf{j}} \frac{1}{\omega^2 - \omega_{\mathbf{j}}^2(\mathbf{k}) - \mathcal{G} \begin{pmatrix} \mathbf{k} \\ \mathbf{j} \end{pmatrix}} \quad (5.2.53)$$

The use of this result together with Eq. (5.2.1) would yield a result for

the Helmholtz free energy of an isotopically disordered cubic Bravais crystal that is exact at least to terms linear in p .

We have not yet said anything about the contour of integration C in Eq. (5.2.1). A convenient choice is the contour shown in Fig. V.5. If we

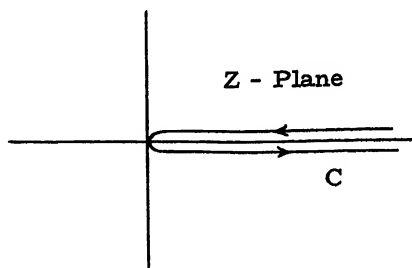


Fig. V.5 A possible contour which can be used in the calculation of the Helmholtz free energy of an isotopically disordered crystal.

use this contour then, as we shall see in the next section, the analysis here is equivalent to solving for the frequency spectrum of the disordered crystal. Alternatively, we can use the contour shown in Fig. III.1. However, since $(d/d\omega) \ln |D(\omega^2)|$ behaves like $|\omega|^{-1}$ as $|\omega| \rightarrow \infty$, we must compute the change in the free energy from its value for the unperturbed mean mass lattice. This is accomplished by replacing $(d/d\omega) \ln |D(\omega^2)|$ in Eq. (5.2.1) by

$$\begin{aligned} & \frac{d}{d\omega} \ln |D(\omega^2)| - \frac{d}{d\omega} \ln |D_0(\omega^2)| \\ &= 2\omega \sum_{kj} \left\{ \frac{1}{\omega^2 - \omega_j^2(k) - \frac{1}{2} \chi_j^2(k)} - \frac{1}{\omega^2 - \omega_j^2(k)} \right\} \end{aligned}$$

The various results obtained in Sec. III can now be used to simplify the calculations.

Finally, we remark that there should be no difficulty, in principle, in extending the calculations of the present section to yield the contribution to $G(kjj')$ which is $O(p^2)$. Such a calculation has been carried out by

Langer for the linear chain.³⁸ The calculation is somewhat more complicated for three-dimensional crystals because of polarization mixing, but in the special case of a cubic Bravais crystal it is likely that the result would take a moderately simple form. Its inclusion in calculations of the free energy would undoubtedly necessitate rather heavy computational effort.

VI. FREQUENCY SPECTRUM OF AN ISOTOPICALLY DISORDERED CRYSTAL

1. Introduction

Although we have shown in the preceding section that the thermodynamic functions of an isotopically disordered crystal can be determined directly, without the necessity of knowing the frequency spectrum of the crystal, the problem of determining the frequency spectrum of a disordered crystal possesses intrinsic interest and a great deal of effort has been expended in attempts to solve this problem.

The first rigorous calculation of the frequency spectrum of a disordered crystal was carried out by Dyson,³⁹ who was concerned with a linear chain with nearest neighbor interactions. He showed that the frequency spectrum can be expressed as a double integral over a function which is the solution to an integral equation. In one special, but nonphysical, case he was able to solve the integral equation and evaluate the double integral, and thus obtain the frequency spectrum. Although Dyson's analysis has subsequently been simplified by Bellman⁴⁰ and by Englman,⁴¹ it has not been possible so far to obtain a result for the frequency spectrum of an isotopically disordered linear chain by his methods. They also have the disadvantage that they are incapable of being extended to crystals of higher dimensions. Schmidt⁴² has also shown very elegantly that the frequency spectrum for the one-dimensional chain can be found in terms of the solution to a certain functional equation. The use of his method is also limited to the one-dimensional case only. The normal mode frequencies of an isotopically disordered linear chain were obtained by second-order perturbation theory by Rubin and Fleishman,⁴³ but no attempt was made to obtain the frequency spectrum.

The first pictorial representations of the frequency spectra of isotopically disordered crystals were obtained by means of the moment trace method.^{23, 44} This method consists of approximating the spectrum by a polynomial of $2n$ -th degree in ω whose coefficients are determined by the conditions that the spectrum yield the even moments of the spectrum

correctly up to μ_{2n} . In the calculations of Domb et al.^{23, 44} the even moments up to μ_{20} were employed. The moment trace method, however, suffers from the fact that the calculations of the higher moments become very tedious and that the convergence of the method may be slow if the spectrum being approximated by a polynomial is not a smoothly varying function of frequency.

A calculation of the frequency spectrum of a disordered linear chain has also been carried out by Hori⁴⁵ via the moment trace method.

Recently Langer³⁸ has calculated the frequency spectrum of a disordered linear chain by a diagrammatic analysis of its perturbation expansion. He obtained an expression for the frequency spectrum which is exact, at least to terms linear in the concentration of the light isotopes. His results show that, because of disorder, an impurity band occurs near the local mode frequency associated with an isolated impurity.

With the development of high-speed computers it has become possible to obtain the frequency spectrum of a long (and finite) disordered linear chain by purely numerical methods. Most of the calculations of this type have been carried out by Dean and his co-workers at the National Physical

Laboratory in England.^{6, 7} The spectra obtained by these authors are characterized by a great deal of fine structure at their high-frequency ends, which has been interpreted by Dean as resulting from localized modes owing to small clusters of light atoms. The theoretical explanation of these results remains one of the more interesting problems in this entire field.

In the present section we describe and give the results of some of the recent work on this problem.

2. The Frequency Spectrum in the Limit of Low Impurity Concentration

In Sec. V we saw that the dynamical matrix for an isotopically disordered crystal is the matrix whose (kj) $(k'j')$ element is given by

$$\begin{aligned} \left(D(w^2) \right)_{\substack{k \\ j} \substack{k' \\ j'}} &= D \left(\begin{matrix} k & k' \\ j & j' \end{matrix}; w^2 \right) = \bar{\Phi} \left(\begin{matrix} k & k' \\ j & j' \end{matrix} \right) \\ &+ \left[w_j^2(k) - w^2 \right] \Delta(k - k') \delta_{jj'} \end{aligned}$$

(6.2.1)

Moreover, we saw that if we denote the eigenvalues of $D(0)$ by w_j^2 ($j=1, 2, \dots, 3N$), then we have the relation

$$\begin{aligned}
 2\omega \sum_j \frac{1}{\omega_j^2 - \omega^2} &= 2\omega \text{Tr } D^{-1}(\omega^2) \\
 &= 2\omega \sum_{kj} D^{-1} \left(\begin{matrix} k & k \\ j & j \end{matrix} ; \omega^2 \right)
 \end{aligned}
 \tag{6.2.2}$$

The frequency spectrum of a crystal is defined by

$$g(\omega) = \frac{2\omega}{3rN} \sum_j \delta(\omega^2 - \omega_j^2) \tag{6.2.3}$$

so that using the fact that the δ function is an even function of its argument we obtain

$$g(\omega) = \frac{2\omega}{3rN} \text{Im} \lim_{\delta \rightarrow 0+} \frac{1}{\pi} \sum_j \frac{1}{\omega_j^2 - \omega^2 - i\delta} \tag{6.2.4}$$

By comparing Eqs. (6.2.2) and (6.2.4), we obtain the result

$$\begin{aligned}
 g(\omega) &= \frac{2\omega}{\pi} \text{Im} \lim_{\delta \rightarrow 0+} \frac{1}{3rN} \text{Tr } D^{-1}(\omega^2 + i\delta) \\
 &= \frac{2\omega}{\pi} \text{Im} \lim_{\delta \rightarrow 0+} \frac{1}{3rN} \sum_{kj} D^{-1} \left(\begin{matrix} k & k \\ j & j \end{matrix} ; \omega^2 + i\delta \right)
 \end{aligned}
 \tag{6.2.5}$$

Again we argue that it is not in the frequency spectrum of a crystal with a particular impurity configuration that we are interested. We can only ask for the frequency spectrum averaged over all possible impurity configurations since we have no a priori knowledge of the impurity configuration in any given crystal. Our basic result, therefore, is

$$\begin{aligned} \langle g(\omega) \rangle_A &= \frac{2\omega}{\pi} \operatorname{Im} \lim_{\delta \rightarrow 0+} \frac{1}{3rN} \\ &\times \sum_{kj} \langle D^{-1} \left(\begin{smallmatrix} k & k \\ j & j \end{smallmatrix}; \omega^2 + i\delta \right) \rangle_A \end{aligned} \quad (6.2.6)$$

We have obtained an expression for $\langle D^{-1}(k j; k' j'; \omega^2) \rangle_A$ for a cubic Bravais crystal in the preceding section exactly to terms linear in the impurity concentration. The result is

$$\begin{aligned} \langle D^{-1} \left(\begin{smallmatrix} k & k' \\ j & j' \end{smallmatrix}; \omega^2 \right) \rangle_A &= \Delta(k - k') \delta_{jj'} \frac{1}{\omega_j^2(k) - \omega^2 + \mathcal{L}(kj)} \\ &= \Delta(k - k') \delta_{jj'} \frac{1}{\omega_j^2(k) - \omega^2 + p\lambda \omega_j^2(k) \frac{1}{1 - \lambda g(\omega^2)}} \end{aligned} \quad (6.2.7)$$

We, therefore, have the formal result that, correct to terms linear in the impurity concentration, the frequency spectrum of this crystal is given by

$$\begin{aligned} \langle g(\omega) \rangle_A &= \frac{2\omega}{\pi} \operatorname{Im} \lim_{\delta \rightarrow 0+} \frac{1}{3N} \\ &\times \sum_{kj} \frac{1}{\omega_j^2(k) - i\delta - \omega^2 + p\lambda \frac{\omega_j^2(k)}{1 - \lambda g(\omega^2 + i\delta)}} \end{aligned} \quad (6.2.8)$$

Because of its simple form we can rewrite Eq. (6.2.8) in a form well suited for numerical evaluation. From Eq. (5.2.49) we obtain

$$\lim_{\delta \rightarrow 0+} g(\omega^2 + i\delta) = -1 + \omega^2 \left[\tilde{G}_0(\omega^2) - i\pi G_0(\omega^2) \right] \quad (6.2.9)$$

where

$$\tilde{G}_0(\omega^2) = \frac{1}{3N} \sum_{kj} \frac{1}{(\omega^2 - \omega_j^2(k))_p} = \int_0^{\omega_L} \frac{g_{oh}(\xi)}{(\omega^2 - \xi^2)_p} d\xi \quad (6.2.10)$$

and where $G_0(\omega^2)$ is the distribution function for the squares of the normal mode frequencies of the unperturbed crystal. In terms of these functions we obtain

$$\langle g(\omega) \rangle_A = \frac{2\omega}{\pi} \operatorname{Im} \lim_{\delta \rightarrow 0+} \frac{1}{3N}$$

$$\times \sum_{kj} \frac{1}{\omega_j^2(k) \left[1 + \frac{\epsilon p}{1 - \epsilon \omega^2 \tilde{G}_0(\omega^2) + i\pi \epsilon \omega^2 G_0(\omega^2)} \right] - \omega^2 - i\delta}$$

We can replace the sums over k and j by an integral over the frequency spectrum of the unperturbed crystal:

$$\langle g(\omega) \rangle_A = \frac{2\omega}{\pi} \operatorname{Im} \lim_{\delta \rightarrow 0+}$$

$$\times \int_0^{\omega_L} \frac{g_0(\xi) d\xi}{\xi^2 \left[1 + \frac{\epsilon p}{1 - \epsilon \omega^2 G_0(\omega^2) + i\pi \epsilon \omega^2 G_0(\omega^2)} \right] - \omega^2 - i\delta}$$

(6.2.11)

It is convenient to introduce the two functions

$$\alpha(\omega) = \frac{1 - \epsilon \omega^2 \tilde{G}_0(\omega^2)}{[1 - \epsilon \omega^2 \tilde{G}_0(\omega^2)]^2 + \pi^2 \epsilon^2 \omega^4 G_0^2(\omega^2)} \quad (6.2.12)$$

$$\beta(\omega) = \frac{\pi \epsilon \omega^2 G_0(\omega^2)}{[1 - \epsilon \omega^2 \tilde{G}_0(\omega^2)]^2 + \pi^2 \epsilon^2 \omega^4 G_0^2(\omega^2)} \quad (6.2.13)$$

Since $G_0(\omega^2)$ is nonvanishing only in the interval $0 < \omega < \omega_L$, $\beta(\omega)$ is nonvanishing in only this interval also. This means that

$$\begin{aligned} \langle g(\omega) \rangle_A &= \frac{2\omega}{\pi} \int_0^{\omega_L} d\xi \, g_0(\xi) \\ &\times \frac{\epsilon p \xi^2 \beta(\omega)}{[\xi^2(1 + \epsilon p \alpha(\omega)) - \omega^2]^2 + \epsilon^2 p^2 \xi^4 \beta^2(\omega)} \quad 0 < \omega < \omega_L \end{aligned} \quad (6.2.14)$$

On the other hand, $G_0(\omega^2)$ vanishes identically for $\omega > \omega_L$. In this case we can write Eq. (6.2.11) as

$$\langle g(\omega) \rangle_A = 2\omega \int_0^{\omega_L} d\xi \, g_0(\xi) \delta[\xi^2(1 + \epsilon p \gamma(\omega)) - \omega^2] \quad \omega > \omega_L$$

where

$$\gamma(\omega) = \frac{1}{1 - \epsilon \omega^2 \tilde{G}_0(\omega^2)}$$
(6.2.16)

Equation (6.2.15) can be rewritten

$$\langle g(\omega) \rangle_A = \frac{g_0 \left(\frac{\omega}{\sqrt{1 + \epsilon p \gamma(\omega)}} \right)}{\sqrt{1 + \epsilon p \gamma(\omega)}} \quad \begin{matrix} \omega > \omega_L \\ 1 + \epsilon p \gamma(\omega) > 0 \end{matrix}$$
(6.2.17)

The preceding results show that a calculation of the frequency spectrum of a disordered cubic Bravais crystal to terms linear in the concentration of light isotopes depends only on our knowing the frequency spectrum of the unperturbed crystal, $g_0(\omega)$. Frequency spectra for many models of crystals have been computed during the past 25 years, and many of these calculations have been carried out for cubic Bravais crystals. Hopefully, it will not be too long before these results are used in the calculation of the frequency spectrum of a disordered crystal. Very recently, Nardelli and Tettamanzi¹⁷ have computed the functions $\tilde{G}_0(\omega^2)$ and $G_0(\omega^2)$ for a nearest neighbor central force model of a face-centered cubic crystal, from Overton and Dent's⁴⁶ results for the frequency spectrum of this lattice. (Note that Nardelli and Tettamanzi normalize their frequency spectra to 3 rather than to 1 as we have done.) Their results could be used to compute the frequency spectrum of a disordered face-centered cubic crystal.

The only evaluation of the expression given by Eq. (6.2.8) up to the present time has been carried out by Langer in the one-dimensional case, where it takes the form

$$\langle g_{1-D}(\omega) \rangle_A = \frac{2\omega}{\pi} \operatorname{Im} \lim_{\delta \rightarrow 0+} \frac{1}{N} \sum_k \frac{1}{\omega_k^2 - i\delta - \omega^2 + p\lambda} \frac{\omega_k^2}{1 - \lambda g(\omega_k^2 + i\delta)}$$
(6.2.18)

since in this case we can suppress the branch or polarization index j . The normal mode frequencies of the unperturbed linear chain have been obtained in Sec. II-1 and are

$$\omega_k^2 = \omega_L^2 \sin^2 \frac{\pi k}{N} \quad k = -\frac{N}{2} + 1, \dots, \frac{N}{2}$$

(6.2.19)

The functions $\tilde{G}_0(\omega^2)$ and $G_0(\omega^2)$ are readily found to be

$$\begin{aligned} \tilde{G}_0(\omega^2) &= 0 & 0 < \omega < \omega_L \\ &= \frac{1}{\omega \sqrt{\omega^2 - \omega_L^2}} & \omega > \omega_L \end{aligned}$$

(6.2.20)

$$\begin{aligned} G_0(\omega^2) &= \frac{1}{\pi} \frac{1}{\omega \sqrt{\omega_L^2 - \omega^2}} & 0 < \omega < \omega_L \\ &= 0 & \omega > \omega_L \end{aligned}$$

(6.2.21)

The integral corresponding to Eq. (6.2.11) is straightforward to evaluate and Langer has expressed the results compactly as

$$\begin{aligned} \langle g_{1-D}(\omega) \rangle_A &= \frac{2}{\pi} \operatorname{Re} \left[\omega_L^2 - \omega^2 \right. \\ &+ \left. \epsilon p \frac{\omega_L^2 \sqrt{\omega_L^2 - \omega^2}}{\sqrt{\omega_L^2 - \omega^2 + i\epsilon\omega}} \right]^{-\frac{1}{2}} \quad 0 < \omega < \omega_L \end{aligned}$$

(6.2.22a)

$$= \frac{2}{\pi} \operatorname{Re} \left[\omega^2 - \omega_L^2 + \epsilon p \frac{\omega_L^2 \sqrt{\omega^2 - \omega_L^2}}{\sqrt{\omega^2 - \omega_L^2 - i\epsilon\omega}} \right]^{-\frac{1}{2}} \quad \omega_L < \omega$$

(6.2.22b)

The spectrum given by Eq. (6.2.22) is plotted in Fig. VI.1 for the particular case $\epsilon = 2/3$, $p = 0.1$.

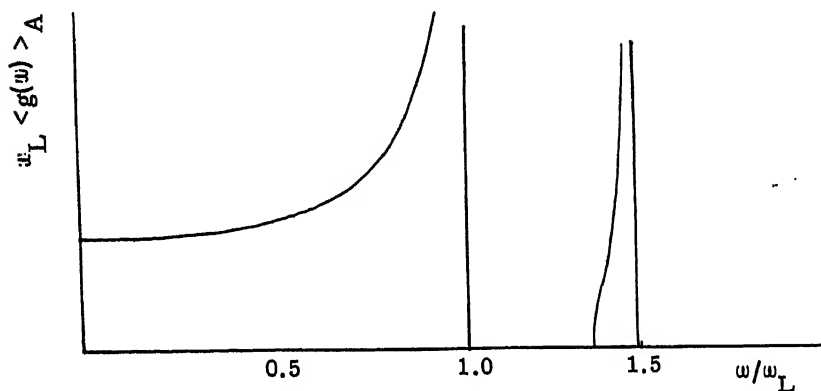


Fig. VI.1 The frequency spectrum of an isotopically disordered linear chain computed by Langer for the case $\epsilon = 2/3$, $p = 0.1$.

The frequency spectrum of a perfect linear chain is found from Eqs. (6.2.21) and (3.1.35) to be

$$\begin{aligned}
 g_0(\omega) &= 2\omega G_0(\omega^2) = \frac{2}{\pi} \frac{1}{\sqrt{\omega_L^2 - \omega^2}} & 0 < \omega < \omega_L \\
 &= 0 & \omega > \omega_L
 \end{aligned}
 \tag{6.2.23}$$

The spectrum calculated by Langer resembles that of the unperturbed crystal in the interval $0 < \omega < \omega_L$, except that the inverse square root singularity at ω_L in the unperturbed spectrum is now replaced by an inverse fourth root singularity. At a frequency slightly above that of the

localized mode associated with an isolated light mass impurity,

$$\omega_{\text{impurity}} = \frac{\omega_L}{\sqrt{1 - \epsilon^2}}$$

an impurity band appears and extends over a small but finite range of frequencies.

Langer's calculation of $\langle g(\omega) \rangle_A$ for a linear chain has recently been improved by Davies.⁴⁷ He started with the expansion for \mathcal{J}_k [the one-dimensional analogue of $\mathcal{J}(kj)$] correct to terms linear in p , which is illustrated diagrammatically in Fig. V.4, and replaced each free phonon propagator d_k^{-1} by the corrected propagator

$$D_k^{-1}(\omega^2) = \frac{1}{\omega_k^2 - \omega^2 + \mathcal{J}_k} \quad (6.2.24)$$

which is the one-dimensional analogue of Eq. (5.2.51). This leads to an implicit equation for the new \mathcal{J}_k since \mathcal{J}_k appears in the modified propagator. The new proper self-energy function, which Davies denotes by $\mathcal{J}_k^{\text{s.c.}}$ (for self-consistent), is then given by the solution of the equation

$$\mathcal{J}_k^{\text{s.c.}}(\omega^2) = p\lambda\omega_k^2 \left\{ 1 + \frac{\lambda}{N} \sum_{k'} \frac{\omega_{k'}^2}{\omega_{k'}^2 - \omega^2 + \mathcal{J}_{k'}^{\text{s.c.}}(\omega^2)} \right\}^{-1} \quad (6.2.25)$$

which replaces the one-dimensional form of Eq. (5.2.50). This equation is readily solved if we make the Ansatz

$$\mathcal{J}_k^{\text{s.c.}}(\omega^2) = A(\omega^2) \omega_k^2$$

(6. 2. 26)

The equation determining $A(\omega^2)$ becomes

$$A(\omega^2) = p\lambda \left\{ 1 + \frac{\lambda}{N} \sum_k \frac{\omega_k^2}{[1 + A(\omega^2)] \omega_k^2 - \omega^2} \right\}^{-1}$$

(6. 2. 27)

In fact one solves for the real and imaginary parts of $A(\omega^2 + i\delta)$ and uses the results to obtain the frequency spectrum from

$$\begin{aligned} \langle g_{1-D}^{\text{s.c.}}(\omega) \rangle_A &= \frac{2\omega}{\pi} \operatorname{Im} \lim_{\delta \rightarrow 0+} \frac{1}{N} \\ &\times \sum_k \frac{1}{\omega_k^2 - i\delta - \omega^2 + \mathcal{J}_k^{\text{s.c.}}(\omega^2 + i\delta)} \end{aligned}$$

(6. 2. 28)

Davies has carried out these calculations, and a plot of the spectrum he has obtained for the case $\epsilon = 2/3$, $p = 0.1$, is shown in Fig. VI. 2. We see that the singular behavior of the impurity band found by Langer is eliminated, and the impurity band is now a smoothly varying function of ω .

There seems to be no difficulty, in principle, in carrying out a similar self-consistent calculation for the frequency spectrum of a disordered cubic Bravais crystal, except that the equation analogous to Eq. (6. 2. 25) will now have to be solved numerically rather than analytically as Davies was able to do.

It is worth mentioning that by using the techniques employed in Sec. V-2 we can carry out a calculation of the inverse matrix

$D^{-1}(kj; k'j'; \omega^2)$ in the case that only one impurity atom is present in the crystal. Without going into the details of the calculation, we remark that the expression for the frequency spectrum obtained in this case

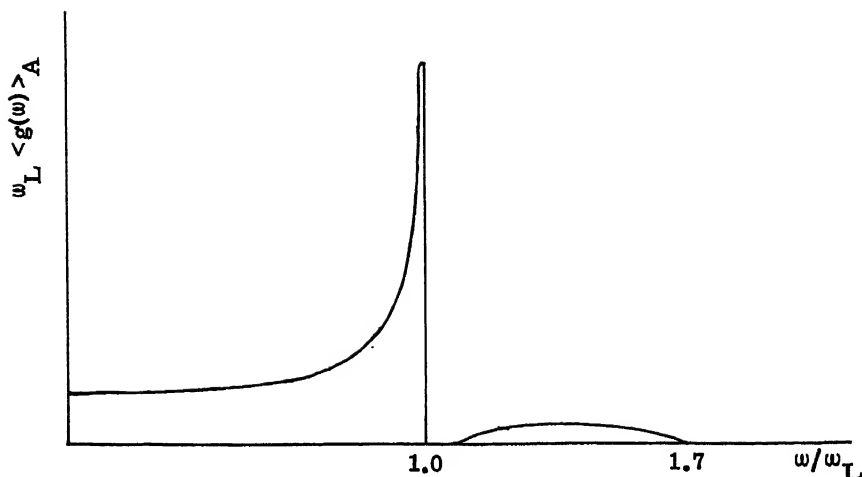


Fig. VI.2 The frequency spectrum of an isotopically disordered linear chain computed in a self-consistent manner by Davis.
 $\epsilon = 2/3$, $p = 0.1$.

is (for a cubic Bravais crystal)

$$\begin{aligned}
 g(\omega) &= \frac{2\omega}{\pi} \operatorname{Im} \lim_{\delta \rightarrow 0+} \frac{1}{3N} \left\{ \sum_{\mathbf{kj}} \frac{1}{\omega^2 - i\delta - \omega_{\mathbf{j}}^2(\mathbf{k})} \right. \\
 &\quad + \frac{\lambda}{N} \sum_{\mathbf{kj}} \frac{\omega_{\mathbf{j}}^2(\mathbf{k})}{(\omega^2 - i\delta - \omega_{\mathbf{j}}^2(\mathbf{k}))^2} \left[1 - \frac{\lambda}{3N} \right. \\
 &\quad \times \left. \sum_{\mathbf{k}_1 \mathbf{j}_1} \frac{\omega_{\mathbf{j}_1}^2(\mathbf{k}_1)}{\omega^2 - i\delta - \omega_{\mathbf{j}_1}^2(\mathbf{k})} \right]^{-1} \left. \right\} \\
 &= g_0(\omega) + \frac{2\omega\epsilon}{N} \\
 &\quad \times \frac{\epsilon\omega^4 [G'_0(\omega^2)\tilde{G}_0(\omega^2) - G_0(\omega^2)\tilde{G}'_0(\omega^2)] - G_0(\omega^2) - \omega^2 G'_0(\omega^2)}{[1 - \epsilon\omega^2 \tilde{G}_0(\omega^2)]^2 + \epsilon^2 \pi^2 \omega^4 G_0^2(\omega^2)} \\
 &\quad 0 < \omega < \omega_L
 \end{aligned}
 \tag{6.2.29a}$$

$$= \frac{1}{N} \delta(\omega - \omega_0) \quad \omega > \omega_L \quad (6.2.29b)$$

In these expressions primes denote differentiation with respect to argument, and ω_0 is the solution of

$$1 = \epsilon \omega^2 \tilde{G}_0(\omega^2) \quad (6.2.30)$$

for $\omega > \omega_L$, and is recognized to be the local mode frequency associated with an isolated light mass defect.

The second term on the right side of Eq. (6.2.29a) can be written compactly as

$$\Delta g(\omega) = \frac{2\omega}{\pi N} \frac{d}{d\omega^2} \tan^{-1} \frac{\pi \epsilon \omega^2 \tilde{G}_0(\omega^2)}{1 - \epsilon \omega^2 \tilde{G}_0(\omega^2)} \quad (6.2.31)$$

This expression holds for $\omega > \omega_L$ as well as for $0 < \omega < \omega_L$, and agrees with a result of Nardelli and Tettamanzi.¹⁷ It is also in agreement with the result obtained by combining Eqs. (3.1.37) and (4.2.1).

The result expressed by Eq. (6.2.31) illustrates a purely computational advantage of working with the function $\Omega(f)$, defined in the present case by Eqs. (4.2.3) and (4.2.4), rather than with $\Delta g(\omega)$ or $\Delta n(\omega)$. This advantage is that the functions $G(f)$ and $G'(f)$ which enter into the expression for $\Omega(f)$ are straightforward to compute numerically and tend to be rather insensitive to the complicated structure of the frequency spectrum of the host crystal. On the other hand, the evaluation of $\tilde{G}_0(\omega^2)$ requires the numerical evaluation of a Cauchy principal value integral, which can be tricky in the case that the integrand is not a smoothly varying function of its argument, and the resulting function is quite sensitive to the singularities in the frequency spectrum.

3. Machine Calculations of the Frequency Spectrum of a Disordered Linear Chain

The inability of analytic calculations of the frequency spectrum of a disordered crystal, such as we have described in Secs. VI-1 and VI-2, to yield more than approximate expressions or expressions valid in more than limiting situations of impurity masses and concentrations, has stimulated attempts to calculate these spectra numerically by essentially brute force methods on high-speed computers. At the present time these calculations have been carried out only for one-dimensional crystals, although an extension of such calculations to two-dimensional lattices has been proposed.

In this section we describe two methods by which frequency spectra of disordered linear chains have been computed numerically on high-speed computers and show examples of spectra computed in this way.

The first of these methods is due to Dean.⁶ His starting point is the equations for the time independent amplitudes $\{u_n\}$ for an isotopically disordered, linear chain with fixed ends, which he writes as

$$\beta_n u_{n-1} + (\alpha_n - \omega^2) u_n + \beta_{n+1} u_{n+1} = 0 \quad n = 1, 2, \dots, N$$

$$u_0 = u_{N+1} = 0 \quad (6.3.1)$$

where

$$\alpha_n = \frac{2\gamma}{M_n} \quad \beta_n = -\frac{\gamma}{\sqrt{M_n M_{n-1}}} \quad (6.3.2)$$

In these equations γ is the nearest neighbor force constant, and M_n is the mass of the n -th atom. Dean then considers the sequence of polynomials $g_0(\omega^2) = 1, g_1(\omega^2), g_2(\omega^2), \dots, g_N(\omega^2)$, where

$$g_i(\omega^2) = \begin{vmatrix} \alpha_1 - \omega^2 & \beta_2 & 0 & \dots & \dots & \dots \\ \beta_2 & \alpha_2 - \omega^2 & \beta_3 & \dots & \dots & \dots \\ 0 & \beta_3 & \alpha_3 - \omega^2 & \dots & \dots & \dots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \beta_i & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \beta_i & \alpha_i - \omega^2 \end{vmatrix} \quad (6.3.3)$$

We are interested in the solutions of $g_N(\omega^2) = 0$.

The successive g 's are related by

$$g_i(\omega^2) = (\alpha_i - \omega^2)g_{i-1}(\omega^2) - \beta_i^2 g_{i-2}(\omega^2) \quad (6.3.4)$$

as is readily verified by expanding the determinant starting with the last column. The polynomials $\{g_i(\omega^2)\}$ thus form a Sturm sequence. A sequence of polynomials $f_1(x), f_2(x), \dots, f_m(x)$ is a Sturm sequence with respect to an interval (a, b) if it has the following two properties:

1. For every value of x ($a < x < b$) if any $f_k(x)$ vanishes, the two adjacent polynomials, $f_{k-1}(x)$ and $f_{k+1}(x)$, are not zero and have opposite signs.
2. The first function in the sequence does not vanish for $a < x < b$.

Since the $\{g_i(\omega^2)\}$ form a Sturm sequence the following theorem holds:

If a and b are real numbers such that $b > a$, then the number of roots of the equation $g_N(\omega^2) = 0$ that lie in the interval (a, b) is equal to $v(b) - v(a)$, where $v(x)$ is the number of variations of sign between consecutive members of the sequence $g_0(x), g_1(x), g_2(x), \dots$.

The cumulative frequency distribution function $N(\omega^2)$, which gives the number of normal modes with squared frequencies less than ω^2 is, therefore, given by

$$N(\omega^2) = v(\omega^2) \quad (6.3.5)$$

In actually performing his numerical calculations, instead of working with the $\{g_i(\omega^2)\}$ Dean found it more convenient to work with the sequence of functions $\{h_i(\omega^2)\}$ defined by

$$h_i(\omega^2) = \frac{g_i(\omega^2)}{g_{i-1}(\omega^2)} \quad (6.3.6)$$

and that satisfy the recurrence relation

$$h_1(w^2) = \alpha_1 - w^2 - \frac{\beta_1^2}{h_1 - 1(w^2)}$$

(6.3.7)

In his earliest numerical calculations of the spectra of isotopically disordered chains, Dean studied diatomic lattices containing equal numbers of atoms of mass M_1 and of mass M_2 . Several random chains were generated of lengths ranging from 2000 atoms to 64,000 atoms. The chains that were actually generated were half the stated length and the second half of any chain was constructed from the first half by replacing all atoms of mass M_1 by atoms of mass M_2 , and vice versa. For a given value of w^2 the functions $h_1(w^2), h_2(w^2), \dots$ were computed, and the number of sign changes which occurred in this sequence were recorded. This number is $v(w^2)$ and hence $N(w^2)$. The distribution function for the squares of the normal mode frequencies, $G(w^2)$, was then obtained from $N(w^2)$ by numerical differentiation. Figure VI.3 shows a typical spectrum obtained by Dean for $M_1/M_2 = 2$.⁶ The dotted line is the spectrum for the ordered ABAB... lattice, the dashed line is the moment trace approximation based on the even moments up to μ_{20} ,²³ and the solid line is the spectrum derived by Dean's Monte Carlo method. Since it is $N(w^2)$ that is calculated rather than $G(w^2) = N'(w^2)$ (which is plotted), there may be some small errors introduced by the process of numerical differentiation.

The work summarized above has been extended to several other cases by Dean and his co-workers.^{6,7,48} In these papers, they have obtained histograms for $G(w^2)$ directly rather than by numerical differentiation of $N(w^2)$. A striking feature of all the spectra obtained by Dean is the finely detailed structure at the higher frequencies. Dean⁶ has interpreted this behavior as resulting from impurity bands whose positions can be calculated approximately by computing the frequencies of various clusters of light atoms in a heavy atom environment.

Dean and Martin⁷ have also published several generalizations of Sturm's theorem that provide a basis for the calculation of the spectra of

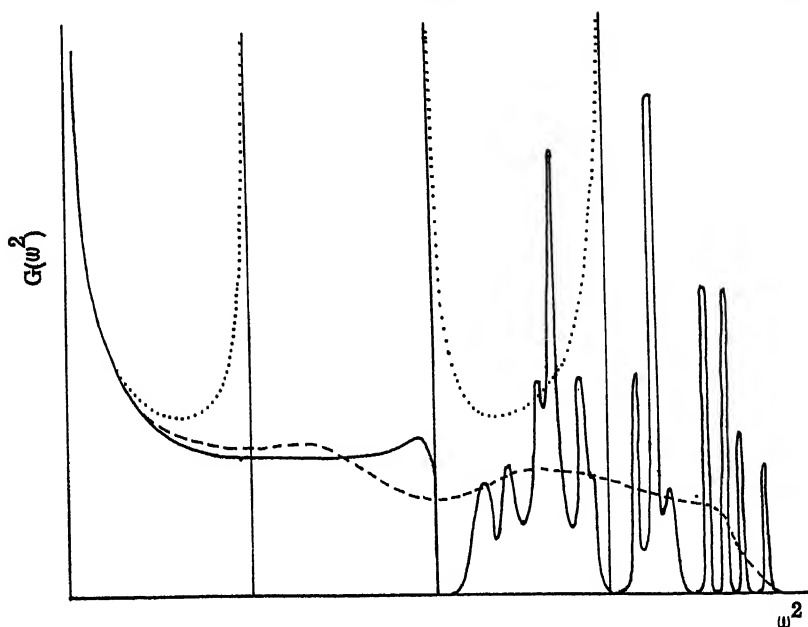


Fig. VI.3 The distribution function for the squares of the normal mode frequencies of an isotopically disordered linear chain for which the mass ratio is 2 and the two constituents are present in equal amounts. The solid line is the result of a machine calculation by Dean⁶ for a linear chain of 32,000 particles. The dashed line is the 20-moment approximation to the spectrum,²³ and the dotted line is the spectrum of an (ordered) alternating, diatomic, linear chain.

two-dimensional lattices and of one-dimensional lattices with longer-range interactions. Martin⁴⁸ has considered the case of a one-dimensional lattice with nearest and next-nearest neighbor interactions. For sufficiently small next-nearest neighbor force constants, he reaches the following qualitative conclusions:

1. The spectrum shows a general drift toward lower frequencies.
2. The peaks at the upper end of the spectrum become higher and narrower.

As these notes were being prepared we learned of a note by Bacon and Dean⁴⁹ in which they present the results of a calculation of the integrated frequency spectrum of a disordered two-dimensional crystal. Their result is shown in Fig. VI.4.

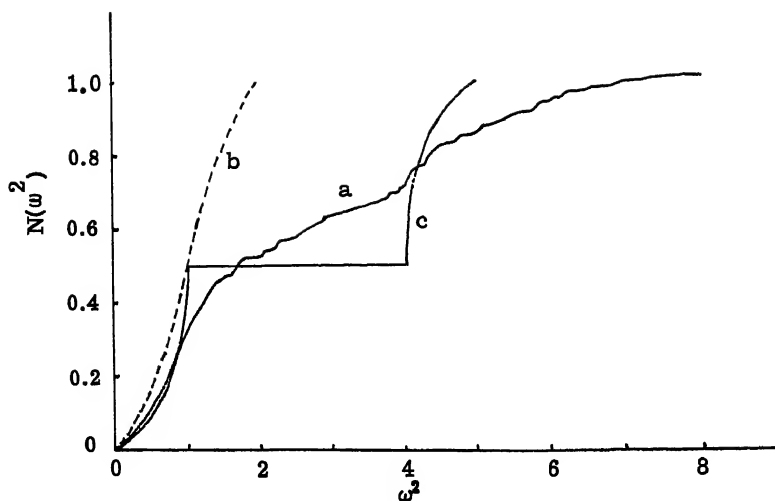


Fig. VI. 4. Integrated frequency spectra for two-dimensional lattices. (a) A disordered lattice in which the ratio of the masses of the two kinds of atoms is 4 to 1, the two constituents are present in equal numbers; (b) the monatomic, heavy mass lattice; (c) the diatomic alternating array.

Recently Rosenstock and McGill⁵⁰ have pointed out an alternative method for the numerical computation of the frequency spectrum of a disordered linear chain, which appears to be somewhat simpler to use in practice than Dean's method. These authors also consider a chain of $N + 2$ atoms with the end atoms fixed. However, they write the equations for the time-independent amplitudes $\{u_n\}$ as

$$\begin{aligned}
 u_0 &= 0 \\
 \gamma u_2 &= (2\gamma - M_1 \omega^2) u_1 \\
 \gamma u_3 &= (2\gamma - M_2 \omega^2) u_2 - \gamma u_1 \\
 \gamma u_4 &= (2\gamma - M_3 \omega^2) u_3 - \gamma u_2 \\
 &\vdots \\
 \gamma u_N &= (2\gamma - M_{N-1} \omega^2) u_{N-1} - \gamma u_{N-2} \\
 \gamma u_{N+1} &= (2\gamma - M_N \omega^2) u_N - \gamma u_{N-1}
 \end{aligned}$$

We can now set $u_1 = 1$ with no loss of generality, since any eigenvalue ω^2 of this set of equations can determine the associated displacement amplitudes only to within a multiplicative constant. Then, for a given

value of ω^2 , we can solve these equations recursively for the remaining $\{u_n\}$. If when we come to u_{N+1} we find that it vanishes, then this value

of ω^2 , is a normal mode frequency of the chain, and the $\{u_n\}$ are the corresponding displacement amplitudes.

The displacement amplitudes $\{u_n\}$ can be regarded as functions of ω^2 , and the zeroes of $u_{N+1}(\omega^2)$ are thus the normal mode frequencies of the chain. Rosenstock and McGill show that the sequence of $\{u_n\}$ computed in this way forms a Sturm sequence. Thus the number of variations in sign that occur if the functions u_1, u_2, \dots, u_{N+1} are successively evaluated for one given value of ω^2 is equal to the number of roots of u_{N+1} that are smaller than ω^2 .

Computationally, then, one constructs a random sequence of N masses of which a certain fraction p are M_1 , and a fraction $1 - p$ are M_2 . The frequency interval $(0, \omega_L^2)$ is then subdivided into some convenient number of intervals $(0, \omega_1^2), (\omega_1^2, \omega_2^2), \dots$. For each ω_j^2 the set of $\{u_n\}$ is computed and the number of sign changes n_j is recorded. The difference $n_j - n_{j-1}$ gives the number of modes with frequencies in the interval $(\omega_{j-1}^2, \omega_j^2)$, and a histogram of the spectrum can be constructed from these data.

Rosenstock and McGill's algorithm has been used by Flinn and Maradudin⁵¹ in calculations of frequency spectra for isotopically disordered linear chains for several of the cases studied by Dean and his co-workers. The results obtained by this method were found to be identical with those obtained by Dean's method, apart from differences that can be ascribed to roundoff.

The problem of reproducing the results of the machine calculations of the frequency spectra of disordered crystals by analytic methods, in particular, reproducing the fine structure displayed by the machine results, remains a challenging one. I think it is intolerable that the solution of such an interesting problem should be left to an electronic computer.

VII. THE PHONON PROPAGATOR

1. The Phonon Propagator for a Disordered Crystal

In this section we carry out the evaluation of a particular displacement correlation function which will play an important role in later applications. This is the function

$$G_{\alpha\beta}(\ell\ell'; u) = \langle Tu_{\alpha}(\ell; u) u_{\beta}(\ell'; 0) \rangle \quad (7.1.1)$$

where $u_{\alpha}(\ell; u)$ is the Heisenberg operator*

$$u_{\alpha}(\ell; u) = e^{uH} u_{\alpha}(\ell; 0) e^{-uH} \quad (7.1.2)$$

The angular brackets in Eq. (7.1.1) denote an average over the canonical ensemble described by the Hamiltonian H , and we have, explicitly,

$$G_{\alpha\beta}(\ell\ell'; u) = \frac{\text{Tr} e^{-\beta H} \left\{ T e^{uH} u_{\alpha}(\ell; 0) e^{-uH} u_{\beta}(\ell'; 0) \right\}}{\text{Tr} e^{-\beta H}} \quad (7.1.3)$$

Here T is the usual time ordering operator that, operating on a product of several time-dependent operators, orders them from right to left in the order of increasing arguments. Thus, for example,

$$\begin{aligned} \langle Tu_{\alpha}(\ell; u) u_{\beta}(\ell'; 0) \rangle &= \langle u_{\alpha}(\ell; u) u_{\beta}(\ell'; 0) \rangle \quad u > 0 \\ &= \langle u_{\beta}(\ell'; 0) u_{\alpha}(\ell; u) \rangle \quad u < 0 \end{aligned}$$

For the applications we have in mind, H is the Hamiltonian for an isotopically disordered crystal. It can be written

*Here and in all that follows we adopt the following convention. The variable t will always denote time, and the Heisenberg operator $O(t)$ is defined by $\exp[i(t/\hbar)H] O(0) \exp[-i(t/\hbar)H]$. For any other variable x the Heisenberg operator $O(x)$ is given by $\exp(xH) O(0) \exp(-xH)$. A similar convention is adopted for operators in the interaction representation.

$$\begin{aligned}
H = & \sum_{\ell \kappa \alpha} \frac{p_{\alpha}^2(\ell)}{2M_{\kappa}} + \frac{1}{2} \sum_{\substack{\ell \kappa \alpha \\ \ell' \kappa' \beta}} \Phi_{\alpha \beta} \left(\frac{\ell \ell'}{\kappa \kappa'} \right) u_{\alpha} \left(\frac{\ell}{\kappa} \right) u_{\beta} \left(\frac{\ell'}{\kappa'} \right) \\
& + \sum_{\ell \kappa \alpha} \left(\frac{1}{2M_{\kappa}^{\ell}} - \frac{1}{2M_{\kappa}} \right) p_{\alpha}^2 \left(\frac{\ell}{\kappa} \right)
\end{aligned}
\tag{7.1.5}$$

where, as usual, M_{κ} denotes a mean mass for the κ -th sublattice. The first two terms on the right of this equation comprise the unperturbed Hamiltonian, while the last term will be treated as the perturbation.

It is now convenient to go over to a representation of the Hamiltonian in terms of phonon creation and destruction operators $a_{\mathbf{k}j}^+$, and $a_{\mathbf{k}j}$, respectively. The relations between the u 's and p 's and the new operators are¹⁶

$$u_{\alpha} \left(\frac{\ell}{\kappa} \right) = \left(\frac{\hbar}{2NM_{\kappa}} \right)^{1/2} \sum_{\mathbf{k}j} \frac{e_{\alpha}(\kappa | \mathbf{k} | j)}{(w_j(\mathbf{k}))^{1/2}} (a_{\mathbf{k}j}^+ - a_{\mathbf{k}j}) e^{2\pi i \mathbf{k} \cdot \mathbf{x}(\ell)}
\tag{7.1.6a}$$

$$p_{\alpha} \left(\frac{\ell}{\kappa} \right) = i \left(\frac{\hbar M_{\kappa}}{2N} \right)^{1/2} \sum_{\mathbf{k}j} e_{\alpha} \left(\kappa | \mathbf{k} | j \right) \left(w_j(\mathbf{k}) \right)^{1/2} (a_{\mathbf{k}j}^+ - a_{\mathbf{k}j}) e^{2\pi i \mathbf{k} \cdot \mathbf{x}(\ell)}
\tag{7.1.6b}$$

The a operators obey the following commutation relations

$$\begin{aligned}
[a_{\mathbf{k}j}, a_{\mathbf{k}'j'}^+] &= \Delta(\mathbf{k} - \mathbf{k}') \delta_{jj'} \\
[a_{\mathbf{k}j}, a_{\mathbf{k}'j'}] &= [a_{\mathbf{k}j}^+, a_{\mathbf{k}'j'}^+] = 0
\end{aligned}
\tag{7.1.7}$$

In terms of these operators the Hamiltonian (7.1.5) becomes (omitting the zero-point energy)

$$H = \sum_{\mathbf{k}\mathbf{j}} \hbar \omega_{\mathbf{j}}(\mathbf{k}) a_{\mathbf{k}\mathbf{j}}^+ a_{\mathbf{k}\mathbf{j}} + \sum_{\substack{\mathbf{k}_1 \mathbf{k}_2 \\ \mathbf{j}_1 \mathbf{j}_2}} V \begin{pmatrix} \mathbf{k}_1 \mathbf{k}_2 \\ \mathbf{j}_1 \mathbf{j}_2 \end{pmatrix} \left[a_{-\mathbf{k}_1 \mathbf{j}_1}^+ a_{\mathbf{k}_2 \mathbf{j}_2} \right. \\ \left. + a_{\mathbf{k}_1 \mathbf{j}_1} a_{-\mathbf{k}_2 \mathbf{j}_2}^+ - a_{-\mathbf{k}_1 \mathbf{j}_1}^+ a_{-\mathbf{k}_2 \mathbf{j}_2} - a_{\mathbf{k}_1 \mathbf{j}_1} a_{\mathbf{k}_2 \mathbf{j}_2} \right] = H_0 + V \quad (7.1.8)$$

$$\left. + a_{\mathbf{k}_1 \mathbf{j}_1} a_{-\mathbf{k}_2 \mathbf{j}_2}^+ - a_{-\mathbf{k}_1 \mathbf{j}_1}^+ a_{-\mathbf{k}_2 \mathbf{j}_2} - a_{\mathbf{k}_1 \mathbf{j}_1} a_{\mathbf{k}_2 \mathbf{j}_2} \right] = H_0 + V \quad (7.1.9)$$

where

$$V \begin{pmatrix} \mathbf{k}_1 \mathbf{k}_2 \\ \mathbf{j}_1 \mathbf{j}_2 \end{pmatrix} = \frac{\hbar}{4N} \sum_{\ell \kappa} \left(\frac{M_{\kappa}}{M_{\ell}^2} - 1 \right) e \left(\kappa \middle| \begin{matrix} \mathbf{k}_1 \\ \mathbf{j}_1 \end{matrix} \right) \cdot e \left(\kappa \middle| \begin{matrix} \mathbf{k}_2 \\ \mathbf{j}_2 \end{matrix} \right) \\ \times \sqrt{\omega_{\mathbf{j}_1}(\mathbf{k}_1) \omega_{\mathbf{j}_2}(\mathbf{k}_2)} e^{2\pi i (\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{x}(\ell)} \quad (7.1.10)$$

Note that $V(\mathbf{k}_1 \mathbf{j}_1; \mathbf{k}_2 \mathbf{j}_2)$ is symmetric in an interchange of $(\mathbf{k}_1 \mathbf{j}_1$ and $(\mathbf{k}_2 \mathbf{j}_2)$.

The correlation function (7.1.1) expressed in terms of the creation and destruction operators becomes

$$G_{\alpha\beta} \left(\ell \ell' ; u \right) = \frac{\hbar}{2N \sqrt{M_{\kappa} M_{\kappa'}}} \sum_{\substack{\mathbf{k} \mathbf{k}' \\ \mathbf{j} \mathbf{j}'}} \frac{e_{\alpha} \left(\kappa \middle| \begin{matrix} \mathbf{k} \\ \mathbf{j} \end{matrix} \right) e_{\beta}^* \left(\kappa' \middle| \begin{matrix} \mathbf{k}' \\ \mathbf{j}' \end{matrix} \right)}{\left(\omega_{\mathbf{j}}(\mathbf{k}) \omega_{\mathbf{j}'}(\mathbf{k}') \right)^{1/2}}$$

$$e^{2\pi i \mathbf{k} \cdot \mathbf{x}(\ell) - 2\pi i \mathbf{k}' \cdot \mathbf{x}(\ell)} F\left(\begin{smallmatrix} \mathbf{k} & \mathbf{k}' \\ \mathbf{j} & \mathbf{j}' \end{smallmatrix}; u\right) \quad (7.1.11)$$

where

$$F\left(\begin{smallmatrix} \mathbf{k} & \mathbf{k}' \\ \mathbf{j} & \mathbf{j}' \end{smallmatrix}; u\right) = \langle T e^{uH} (a_{-\mathbf{k}\mathbf{j}}^+ + a_{\mathbf{k}\mathbf{j}}) e^{-uH} (a_{-\mathbf{k}'\mathbf{j}'} + a_{\mathbf{k}'\mathbf{j}'}^+) \rangle \quad (7.1.12)$$

If we expand $F(\mathbf{k}\mathbf{j}; \mathbf{k}'\mathbf{j}'; u)$ in terms of the eigenstates of H ,

$$F\left(\begin{smallmatrix} \mathbf{k} & \mathbf{k}' \\ \mathbf{j} & \mathbf{j}' \end{smallmatrix}; u\right) = \frac{1}{Z} \sum_{mn} e^{-\beta E_m} e^{u(E_m - E_n)} \times \langle m | a_{-\mathbf{k}\mathbf{j}}^+ + a_{\mathbf{k}\mathbf{j}} | n \rangle \langle n | a_{-\mathbf{k}'\mathbf{j}'} + a_{\mathbf{k}'\mathbf{j}'}^+ | m \rangle \quad u > 0 \quad (7.1.13a)$$

$$= \frac{1}{Z} \sum_{mn} e^{-\beta E_m} e^{-u(E_m - E_n)}$$

$$\times \langle m | a_{-\mathbf{k}'\mathbf{j}'} + a_{\mathbf{k}'\mathbf{j}'}^+ | n \rangle \langle n | a_{-\mathbf{k}\mathbf{j}} + a_{\mathbf{k}\mathbf{j}}^+ | m \rangle \quad u < 0 \quad (7.1.13b)$$

where Z is the crystal's partition function, then we see that it satisfies the following periodicity condition:

$$F\left(\begin{smallmatrix} \mathbf{k} & \mathbf{k}' \\ \mathbf{j} & \mathbf{j}' \end{smallmatrix}; u + \beta\right) = F\left(\begin{smallmatrix} \mathbf{k} & \mathbf{k}' \\ \mathbf{j} & \mathbf{j}' \end{smallmatrix}; u\right) \quad -\beta < u < 0 \quad (7.1.14)$$

This means that if we know $F(kj; k'j'; u)$ for u in the interval $(-\beta, 0)$ we know it in the entire interval $(-\beta, \beta)$. The form of Eq. (7.1.14) suggests that we expand $F(kj; k'j'; u)$ in a Fourier series with a period β :

$$F\left(\begin{matrix} k & k' \\ j & j' \end{matrix}; u\right) = \sum_{\ell=-\infty}^{\infty} A_{\ell}\left(\begin{matrix} k & k' \\ j & j' \end{matrix}\right) e^{\frac{2\pi i \ell u}{\beta}} \quad (7.1.15)$$

The Fourier coefficient $A_{\ell}\left(\begin{matrix} k & k' \\ j & j' \end{matrix}\right)$ is given by

$$A_{\ell}\left(\begin{matrix} k & k' \\ j & j' \end{matrix}\right) = \frac{1}{\beta} \int_0^{\beta} F\left(\begin{matrix} k & k' \\ j & j' \end{matrix}; u\right) e^{-\frac{2\pi i \ell u}{\beta}} du \quad (7.1.16)$$

In the application of the results that we obtain in this section, we shall see that it is the Fourier coefficients $\{A_{\ell}\}$ that are ultimately required. It is therefore with their evaluation that we will be primarily concerned in what follows.

Our starting point for the calculation of $F(kj; k'j'; u)$ and hence of $A_{\ell}(kj; k'j')$ is the expansion

$$\begin{aligned} F\left(\begin{matrix} k & k' \\ j & j' \end{matrix}; u\right) &= \langle T(\tilde{a}_{-kj}^{+}(u) + \tilde{a}_{kj}^{-}(u))(\tilde{a}_{-k'j'}(0) + \tilde{a}_{k'j'}^{+}(0)) \\ &\times \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int_0^{\beta} d\beta_1 \dots \int_0^{\beta} d\beta_n \tilde{V}(\beta_1) \dots \tilde{V}(\beta_n) \rangle_{oc} \end{aligned} \quad (7.1.17)$$

which has been derived elsewhere.⁵² In this expansion $\tilde{O}(x)$ is an operator in the interaction representation and is defined by

$$\tilde{O}(x) = e^{xH_0} O(0) e^{-xH_0} \quad (7.1.18)$$

where H_0 is the unperturbed Hamiltonian. In particular we have the results

$$e^{xH_0} a_{kj} e^{-xH_0} = a_{kj} e^{-x\hbar\omega_j(k)}$$

$$e^{xH_0} a_{kj}^+ e^{-xH_0} = a_{kj}^+ e^{x\hbar\omega_j(k)}$$
(7.1.19)

The suffix 0 on the brackets in Eq. (7.1.17) means that the thermal average is to be carried out over the canonical ensemble of the unperturbed system, that is, with respect to the weight function

$$\rho_0 = \frac{e^{-\beta H_0}}{Z_0}$$
(7.1.20)

where Z_0 is the partition function for the unperturbed crystal. The suffix c attached to the same brackets means that when we draw diagrams to represent the terms in the expansion (7.1.17) we must consider only connected diagrams. Any diagram made up of vertices with lines joining them is called connected if we can get from any vertex to any other vertex by staying on the lines of the diagram. We shall return to this point in a moment.

If we recall the form of the perturbation Hamiltonian V , Eq. (7.1.8), we see that Eq. (7.1.17) is the sum of the thermal averages of products of phonon creation and destruction operators. If we work in the representation of the eigenstates of H , which we can do since a trace is invariant against representation, then ρ_0 is a diagonal operator. This means that,

in evaluating the thermal average of a product of creation and destruction operators, we obtain a nonvanishing contribution only from products in which the number of creation operators associated with a mode (kj) equals the number of destruction operators associated with the same mode.

The only nonvanishing averages of a product of two operators are

$$\begin{aligned}
 \langle a_{kj}^+ a_{kj} \rangle_0 &= \frac{\sum_{n=0}^{\infty} n e^{-n\beta \hbar \omega_j(k)}}{\sum_{n=0}^{\infty} e^{-n\beta \hbar \omega_j(k)}} \\
 &= n_j(k) = \frac{1}{e^{\beta \hbar \omega_j(k)} - 1}
 \end{aligned}
 \tag{7.1.21a}$$

$$\langle a_{kj} a_{kj}^+ \rangle_0 = n_j(k) + 1 = \frac{1}{1 - e^{-\beta \hbar \omega_j(k)}}
 \tag{7.1.21b}$$

$n_j(k)$ is the mean phonon occupation number. The averages (7.1.21) are called contractions.

Since ρ_0 is a product of operators, one for each mode,

$$\rho_0 = \prod_{kj} \frac{e^{-\beta \hbar \omega_j(k) a_{kj}^+ a_{kj}}}{1 - e^{-\beta \hbar \omega_j(k)}}$$

then the average of a product of creation and destruction operators is just the product of the averages of the products of the operators belonging to the same mode, e.g.,

$$\begin{aligned}
 &\langle a_{q_1}^+ a_{q_2} a_{q_3}^+ a_{q_1} a_{q_3} a_{q_2}^+ \rangle_0 \\
 &= \langle a_{q_1}^+ a_{q_1} \rangle_0 \langle a_{q_2} a_{q_2}^+ \rangle_0 \langle a_{q_3}^+ a_{q_3} \rangle_0
 \end{aligned}$$

where we have used q to denote the index pair (kj) .

However, it can happen that for a certain mode (kj) we can have more than one creation and destruction operator present. The average of such a product is obtained from an extension of Wick's theorem due to Bloch and DeDominicis⁵³; it is the sum of the products of all the averages obtained by pairing the creation and destruction operators corresponding to the same mode two-by-two different in all possible ways. For example we have

$$\begin{aligned}
\langle a_q^+ a_q a_q^+ a_q \rangle_0 &= \langle a_q^+ a_q \rangle_0 \langle a_q^+ a_q \rangle_0 + \langle a_q^+ a_q \rangle_0 \langle a_q^+ a_q \rangle_0 \\
&= n_q^2 + n_q (n_q + 1) \\
&= n_q (2n_q + 1)
\end{aligned}$$

This result means that in evaluating the thermal averages that appear in the expansion (7.1.17) we need only pair the creation and destruction operators two-by-two different in all possible ways which give rise to connected diagrams, and evaluate the thermal averages with the aid of Eqs. (7.1.19) and (7.1.21).

The various terms in the expansion (7.1.17) can be represented by diagrams. The vertices of each diagram correspond to the interactions described by the $\{V(\beta)\}$ and are represented by dots. In addition, the points u and o , at which a phonon is emitted or absorbed, are also considered to be (external) vertices. We order the vertices from the bottom of the diagram in the order of increasing "times" $0, \beta_n,$

$\beta_{n-1}, \dots, \beta_1, u$. Each contraction is represented by a solid line joining two vertices, and the lines are directed. The contraction

$\langle a_q^+ a_q \rangle_0$ is represented by a line directed downward, whereas the contraction $\langle a_q a_q^+ \rangle_0$ is represented by a line directed upward. This

corresponds to associating a creation operator with a line entering a vertex and a destruction operator with a line leaving a vertex. Because of the restriction to connected diagrams and the fact that the interaction vertices describe two phonon processes, whereas the vertices at u and 0 correspond to single phonon processes, the only kinds of diagrams we can draw are those in which a continuous line joins the vertices at u and 0 and threads each of the interaction vertices once. The zero-order and first-order diagrams contributing to $F(kj; k'j'; u)$ are shown in Figs. VII.1a and VII.1b, respectively. With each interaction vertex we associate a "matrix element" $V(k_1 j_1; k_2 j_2)$. However, we have to note

that because of the form of the perturbation Hamiltonian V , Eq. (7.1.8), some vertices have a "plus" sign associated with this matrix element, others have a "minus" sign. The rule for affixing the correct sign to the vertex is that vertices with one line entering the vertex and one line leaving it have a "plus" sign attached, while a vertex with both lines directed into it or both lines directed away from it, has a "minus" sign attached.

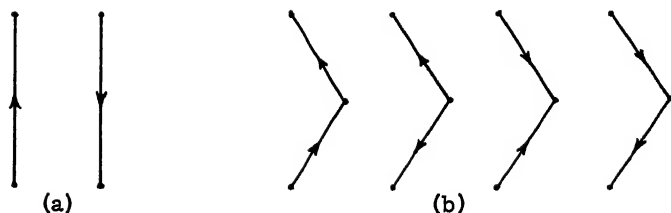


Fig. VII. 1. (a) The zero-order diagrams which contribute to $F(kj; k'j'; u)$. (b) The first-order diagrams which contribute to $F(kj; k'j'; u)$.

Before proceeding to the evaluation of $F(kj; k'j'; u)$ we bring out one more point. In an n -th order connected diagram there appear $n + 1$ lines corresponding to the propagation of a free phonon. This means that there are 2^{n+1} possible diagrams for a fixed order of the vertices corresponding to the fact that each of these $n + 1$ lines can be directed either up or down. This fact has been illustrated in Fig. VII. 1 for the zero-order and first-order diagrams. Let us now focus our attention on one of these 2^{n+1} n -th order diagrams, say the second-order diagram shown in Fig. VII. 2a. For the particular order of the vertices shown in Fig. VII. 2, there exists a second diagram, shown in Fig. VII. 2b.

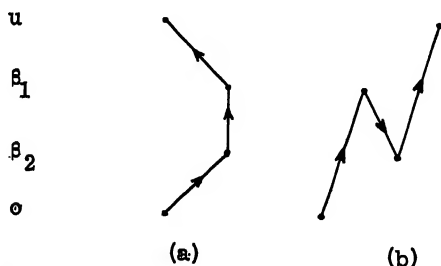


Fig. VII. 2. Two topologically equivalent diagrams which contribute in second order to $F(kj; k'j'; u)$.

However, if we calculate the contribution to $F(kj; k'j'; u)$ associated with the diagram in Fig. VII. 2b, we find that it is equal to the contribution associated with the diagram in Fig. VII. 2a. The reason is that the expressions for the two contributions differ only in the interchange of two dummy integration variables. This result can be generalized to n -th order.

In n -th order a simple combinatorial argument shows that for a given order of vertices, for each one of the $2n+1$ possible choices for the distributions of directions for the free phonon lines, we can draw

$n!$ diagrams, which correspond to this number of ways of joining the n interaction vertices and two external vertices to yield a connected diagram.

Each of these $n!$ diagrams contributes equally to $F(k_j; k'_j; u)$. This means that we need to evaluate the contribution associated with only one of these diagrams and multiply it by $n!$. We can characterize these $n!$ diagrams by saying that they are topologically equivalent, that is, we can obtain any one of them from any one of the remaining diagrams by continuously deforming the latter.

The operations of taking the thermal average and integrating over the $\{\beta\}$ variables in Eq. (7.1.17) can be considerably simplified if we introduce the following function:

$$\begin{aligned} A_{kj}(u_1, u_2) &= \langle T a_{kj}^+(u_1) a_{kj}(u_2) \rangle_0 \\ &= n_j(k) e^{(u_1 - u_2) \hbar \omega_j(k)} \quad u_1 > u_2 \\ &= (n_j(k) + 1) e^{(u_1 - u_2) \hbar \omega_j(k)} \quad u_1 < u_2 \end{aligned}$$

This function has a simple Fourier series expansion

$$\begin{aligned} A_{kj}(u) &= \sum_{\ell=-\infty}^{\infty} a_{\ell} \left(\frac{k}{j} \right) e^{\frac{2\pi i \ell u}{\beta}} \\ &= \frac{1}{\beta \hbar} \sum_{\ell=-\infty}^{\infty} \frac{e^{\frac{2\pi i \ell u}{\beta}}}{\omega_j(k) - i\omega_{\ell}} \end{aligned}$$

(7.1.23)

where $\omega_{\ell} = 2\pi\ell / \beta\hbar$. In the rest of this section we shall call

$a_{\ell}(k)$ the free phonon propagator.

Although on the face of it it may appear that $A_{kj}(u_1, u_2)$ can only be used in a diagram in connection with lines directed downward, from the fact that

$$\langle Ta_{kj}^+(u_1) a_{kj}(u_2) \rangle_0 = \langle Ta_{kj}(u_2) a_{kj}^+(u_1) \rangle_0 \quad (7.1.24)$$

we can devise a convention which permits the application of Eq. (7.1.23) to either a line directed upward or to a line directed downward. This convention is that, for a line joining two vertices at times u_1 and u_2 , we write the exponential factor in Eq. (7.1.23) as $\exp \{ \pm (2\pi i \ell / \beta) \times (u_1 - u_2) \}$, where the sign is determined in such a way that ℓ for the line leaving a vertex is prefixed by a minus sign and ℓ for the line entering a vertex is prefixed by a plus sign.

If the Fourier series (7.1.23) is substituted into the expansion (7.1.17) for any particular diagram, the integrals over the β 's can be carried out unrestrictedly over the interval $(0, \beta)$ since the manner in which the series is constructed ensures that the time orderings are taken into account properly. The following result is of use in carrying out these integrations:

$$\int_0^\beta e^{\frac{2\pi i n u}{\beta}} du = \beta \delta_{n,0} \quad (7.1.25)$$

where n is an integer. The fact that there are only two exponential factors associated with each interaction vertex, together with the result (7.1.25), implies a conservation condition for the ℓ 's at a vertex, namely, the sum of the ℓ 's coming into a vertex must equal the sum of the ℓ 's leaving the vertex. The result of the integrations is a function which depends on u in the form $\exp \{ \pm 2\pi i \ell u / \beta \}$, so that its contribution to $A_{\ell}(k; k'j)$ can be read off simply.

However, if we are interested in formulating a systematic procedure for obtaining $A_{\ell}(k; k'j)$ directly, we must have a way of ensuring

that the function of u which results after the integrations over the β variables are carried out is of the form $\exp \{ +2\pi i \ell u / \beta \}$ rather than of the form $\exp \{ -2\pi i \ell u / \beta \}$. This is so that we do not have to keep track of the various $\exp \{ 2\pi i \ell \beta / \beta \}$ factors associated with each vertex, and do not have to worry about changing the dummy summation variable from ℓ to $-\ell$ if our result comes out proportional to

$\exp \{-2\pi i \ell u / \beta\}$. We can accomplish this end if we combine the convention for writing the exponential factors in Eq. (7.1.23) described two paragraphs above, with the conservation condition on the ω_ℓ 's stated in the preceding paragraph. The rule is simple. With a line directed upward associate a propagator $a_{+\ell}(k_{\ell j})$; with a line directed downward associate a propagator $a_{-\ell}(k_{\ell j})$. The integrations over the β variables are then accounted for by the conservation condition on the ω_ℓ 's at each vertex, together with a factor of β^n in n-th order, which is a consequence of Eq. (7.1.25). For convenience we can always associate the propagator $a_{\pm\ell}(k'j')$ with the line joining the vertices at β_1 and u . We shall, then, always automatically obtain the n-th order contribution to $A_{\ell}(k j; k' j')$.

We can now summarize the rules for computing the n-th order contribution to the Fourier coefficient $A_{\ell}(k j; k' j')$.

1. Draw all topologically distinct connected n-th order diagrams in which a free phonon line labeled by $(k j)$ enters or leaves the diagram at the bottom of the page and a free phonon line labeled by $(k' j')$ leaves or enters the diagram at the top of the page.

2. With each free phonon line labeled $(k j)$ and directed upward associate a factor

$$a_{+\ell}\left(\begin{smallmatrix} k \\ j \end{smallmatrix}\right) = \frac{1}{\beta \hbar} \frac{1}{\omega_j(k) - i\omega_\ell} \quad \omega_\ell = \frac{2\pi\ell}{\beta \hbar}$$

With each free phonon line labeled $(k j)$ and directed downward associate a factor

$$a_{-\ell}\left(\begin{smallmatrix} k \\ j \end{smallmatrix}\right) = \frac{1}{\beta \hbar} \frac{1}{\omega_j(k) + i\omega_\ell} \quad \omega_\ell = \frac{2\pi\ell}{\beta \hbar}$$

3. At each interaction vertex conserve the ω_ℓ 's according to the rule that the sum of the ω_ℓ 's leaving a vertex equals the sum of the ω_ℓ 's entering the vertex.

4. At each vertex insert the appropriate matrix element.

5. Insert a factor $(-1)^n \beta^n / n!$, where the β^n comes from the integrations over the $n\beta$ variables.

6. Insert a combinatorial factor, which is $(-1)^{n_d} n! 2^n$. n_d is the number of vertices in a diagram that have both free phonon lines directed into it or both lines directed out of it. The factor $(-1)^{n_d}$ takes account of the fact that some of the matrix elements in the perturbation Hamiltonian are negative. The factor of $n!$ takes care of the $n!$ topologically equivalent diagrams for a particular choice of the directions of the free phonon lines. The final factor of 2^n comes from the fact that there are always two phonons at a given vertex with which an incoming (or outgoing) phonon line can pair, and that each of these two possible pairings gives the same contribution to $A_\ell(kj; k'j')$ because $V(k_1j_1; k_2j_2)$ is symmetric in (k_1j_1) and (k_2j_2) .

7. Finally, sum over the independent k 's, j 's, and ℓ 's.

We now work out explicitly the first few terms in the expansion of $A_\ell(kj; k'j')$. In zeroth-order we obtain as the contributions associated with the diagram of Fig. VII. 1a,

$$A_\ell^{(0)} \begin{pmatrix} k & k' \\ j & j' \end{pmatrix} = \left[a_{\ell} \begin{pmatrix} k \\ j \end{pmatrix} + a_{-\ell} \begin{pmatrix} k \\ j \end{pmatrix} \right] \Delta(k-k') \delta_{jj'} \quad (7.1.26)$$

In first-order, diagrams of Fig. VII. 1b contribute

$$A_\ell^{(1)} \begin{pmatrix} k & k' \\ j & j' \end{pmatrix} = \frac{(-1)^1}{1!} \beta(1!) 2 \left\{ a_{\ell} \begin{pmatrix} k \\ j \end{pmatrix} V \begin{pmatrix} -k & k' \\ j & j' \end{pmatrix} a_{\ell} \begin{pmatrix} k' \\ j' \end{pmatrix} \right. \\ \left. - a_{-\ell} \begin{pmatrix} k \\ j \end{pmatrix} V \begin{pmatrix} -k & k' \\ j & j' \end{pmatrix} a_{-\ell} \begin{pmatrix} k' \\ j' \end{pmatrix} \right\}$$

$$\begin{aligned}
& - a_{\ell(j)} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) V \left(\begin{smallmatrix} -k & k' \\ j & j' \end{smallmatrix} \right) a_{-\ell(j')} \left(\begin{smallmatrix} k' \\ j' \end{smallmatrix} \right) \\
& + a_{-\ell(j)} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) V \left(\begin{smallmatrix} -k & k' \\ j & j' \end{smallmatrix} \right) a_{-\ell(j')} \left(\begin{smallmatrix} k' \\ j' \end{smallmatrix} \right) \Big\} \\
& = -2\beta \left[a_{\ell(j)} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) - a_{-\ell(j)} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) \right] V \left(\begin{smallmatrix} -k & k' \\ j & j' \end{smallmatrix} \right) \left[a_{\ell(j')} \left(\begin{smallmatrix} k' \\ j' \end{smallmatrix} \right) - a_{-\ell(j')} \left(\begin{smallmatrix} k' \\ j' \end{smallmatrix} \right) \right]
\end{aligned}
\tag{7.1.27}$$

That the matrix element appearing in this expression has one of the wave vectors k_j set equal to $-k$ is due to the fact that to obtain a non-vanishing contraction between the operators a_{-kj}^+ and $a_{k'j'}$, k' has to equal $-k$ (and j' has to equal j).

To simplify the appearance of subsequent expressions let us introduce the abbreviations

$$\alpha_{\ell(j)} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) = a_{\ell(j)} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) + a_{-\ell(j)} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) = \frac{2\omega_j(k)}{\beta \hbar} \frac{1}{\omega_j^2(k) + \omega_{\ell}^2}
\tag{7.1.28a}$$

$$\beta_{\ell(j)} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) = a_{\ell(j)} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) - a_{-\ell(j)} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) = \frac{2i\omega_{\ell}}{\beta \hbar} \frac{1}{\omega_j^2(k) + \omega_{\ell}^2}
\tag{7.1.28b}$$

In terms of these coefficients our previous results become

$$A_{\ell}^{(0)} \left(\begin{smallmatrix} k & k' \\ j & j' \end{smallmatrix} \right) = \alpha_{\ell(j)} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) \Delta(k - k') \delta_{jj'}
\tag{7.1.29a}$$

$$A_{\ell}^{(1)} \left(\begin{smallmatrix} k & k' \\ j & j' \end{smallmatrix} \right) = -2\beta \beta_{\ell(j)} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) V \left(\begin{smallmatrix} -k & k' \\ j & j' \end{smallmatrix} \right) \beta_{\ell(j')} \left(\begin{smallmatrix} k' \\ j' \end{smallmatrix} \right)
\tag{7.1.29b}$$

If we go to second order, we find that

$$\begin{aligned}
A_{\ell}^{(2)} \left(\begin{matrix} k & k' \\ j & j' \end{matrix} \right) &= \frac{(-1)^2}{2!} \beta^2 2! 2^2 \sum_{\substack{k_1 \\ j_1}} \left\{ a_{\ell} \left(\begin{matrix} k \\ j \end{matrix} \right) V \left(\begin{matrix} -k & k_1 \\ j & j_1 \end{matrix} \right) a_{\ell} \left(\begin{matrix} k_1 \\ j_1 \end{matrix} \right) V \left(\begin{matrix} -k_1 & k' \\ j_1 & j' \end{matrix} \right) a_{\ell} \left(\begin{matrix} k' \\ j' \end{matrix} \right) \right. \\
&\quad - a_{-\ell} \left(\begin{matrix} k \\ j \end{matrix} \right) V \left(\begin{matrix} -k & k_1 \\ j & j_1 \end{matrix} \right) a_{\ell} \left(\begin{matrix} k_1 \\ j_1 \end{matrix} \right) V \left(\begin{matrix} -k_1 & k' \\ j_1 & j' \end{matrix} \right) a_{\ell} \left(\begin{matrix} k' \\ j' \end{matrix} \right) \\
&\quad + a_{\ell} \left(\begin{matrix} k \\ j \end{matrix} \right) V \left(\begin{matrix} -k & k_1 \\ j & j_1 \end{matrix} \right) a_{-\ell} \left(\begin{matrix} k_1 \\ j_1 \end{matrix} \right) V \left(\begin{matrix} -k_1 & k' \\ j_1 & j' \end{matrix} \right) a_{\ell} \left(\begin{matrix} k' \\ j' \end{matrix} \right) \\
&\quad - a_{-\ell} \left(\begin{matrix} k \\ j \end{matrix} \right) V \left(\begin{matrix} -k & k_1 \\ j & j_1 \end{matrix} \right) a_{-\ell} \left(\begin{matrix} k_1 \\ j_1 \end{matrix} \right) V \left(\begin{matrix} -k_1 & k' \\ j_1 & j' \end{matrix} \right) a_{\ell} \left(\begin{matrix} k' \\ j' \end{matrix} \right) \\
&\quad - a_{\ell} \left(\begin{matrix} k \\ j \end{matrix} \right) V \left(\begin{matrix} -k & k_1 \\ j & j_1 \end{matrix} \right) a_{\ell} \left(\begin{matrix} k_1 \\ j_1 \end{matrix} \right) V \left(\begin{matrix} -k_1 & k' \\ j_1 & j' \end{matrix} \right) a_{-\ell} \left(\begin{matrix} k' \\ j' \end{matrix} \right) \\
&\quad + a_{-\ell} \left(\begin{matrix} k \\ j \end{matrix} \right) V \left(\begin{matrix} -k & k_1 \\ j & j_1 \end{matrix} \right) a_{\ell} \left(\begin{matrix} k_1 \\ j_1 \end{matrix} \right) V \left(\begin{matrix} -k_1 & k' \\ j_1 & j' \end{matrix} \right) a_{-\ell} \left(\begin{matrix} k' \\ j' \end{matrix} \right) \\
&\quad - a_{\ell} \left(\begin{matrix} k \\ j \end{matrix} \right) V \left(\begin{matrix} -k & k_1 \\ j & j_1 \end{matrix} \right) a_{-\ell} \left(\begin{matrix} k_1 \\ j_1 \end{matrix} \right) V \left(\begin{matrix} -k_1 & k' \\ j_1 & j' \end{matrix} \right) a_{-\ell} \left(\begin{matrix} k' \\ j' \end{matrix} \right) \\
&\quad \left. + a_{-\ell} \left(\begin{matrix} k \\ j \end{matrix} \right) V \left(\begin{matrix} -k & k_1 \\ j & j_1 \end{matrix} \right) a_{-\ell} \left(\begin{matrix} k_1 \\ j_1 \end{matrix} \right) V \left(\begin{matrix} -k_1 & k' \\ j_1 & j' \end{matrix} \right) a_{-\ell} \left(\begin{matrix} k' \\ j' \end{matrix} \right) \right\} \\
&= (2\beta)^2 \sum_{k_1 j_1} \beta_{\ell} \left(\begin{matrix} k \\ j \end{matrix} \right) V \left(\begin{matrix} -k & k_1 \\ j & j_1 \end{matrix} \right) \alpha_{\ell} \left(\begin{matrix} k_1 \\ j_1 \end{matrix} \right) V \left(\begin{matrix} -k_1 & k' \\ j_1 & j' \end{matrix} \right) \beta_{\ell} \left(\begin{matrix} k' \\ j' \end{matrix} \right)
\end{aligned}
\tag{7.1.30}$$

The succeeding terms in the expansion for $A_{\ell}(kj; k'j')$ have the same structure. Thus we can write

$$\begin{aligned}
A_{\ell(jj')}^{(kk')} &= \alpha_{\ell(j)}^{(k)} \Delta(k-k') \delta_{jj'} - \beta_{\ell(j)}^{(k)} \left\{ 2\beta V \begin{pmatrix} -k & k' \\ j & j' \end{pmatrix} \right. \\
&\quad - (2\beta)^2 \sum_{\substack{k_1 \\ j_1}} V \begin{pmatrix} -k & k_1 \\ j & j_1 \end{pmatrix} \alpha_{\ell(j_1)}^{(k_1)} V \begin{pmatrix} k_1 & k' \\ j_1 & j' \end{pmatrix} \\
&\quad + (2\beta)^3 \sum_{\substack{k_1 k_2 \\ j_1 j_2}} V \begin{pmatrix} -k & k_1 \\ j & j_1 \end{pmatrix} \alpha_{\ell(j_1)}^{(k_1)} V \begin{pmatrix} -k_1 & k_2 \\ j_1 & j_2 \end{pmatrix} \alpha_{\ell(j_2)}^{(k_2)} V \begin{pmatrix} -k_2 & k' \\ j_2 & j' \end{pmatrix} \\
&\quad \left. - \dots \right\} \beta_{\ell(j')}^{(k')} \quad (7.1.31)
\end{aligned}$$

On comparing Eqs. (7.1.31) and (5.2.9) we see a strong formal similarity between the two expansions. Indeed the coefficient $V \begin{pmatrix} -k & k' \\ j & j' \end{pmatrix}$

differs from the coefficient $\beta(kj'; k'j')$ defined by Eq. (5.1.6)

only in the replacement of the factor $\omega_{j'}^2(k')$ appearing in the latter

function by $\chi/4 \sqrt{\omega_j(k) \omega_{j'}(k')}$. Thus we can use the results

of Sec. V-2 to complete the analysis of this section.

In applications of the results of this section we shall not be interested so much in $A_{\ell}(kj; k'j')$ but in its value averaged over all configurations of impurities. We know now that this average is diagonal in k and k' and we can write it as

$$\begin{aligned}
\langle A_{\ell(jj')}^{(kk')} \rangle_A &= \Delta(k-k') A_{\ell(jj')}^{(k)} = \Delta(k-k') \left\{ \delta_{jj'} \alpha_{\ell(j)}^{(k)} \right. \\
&\quad - \beta_{\ell(j)}^{(k)} \left[2\beta \langle V \begin{pmatrix} -k & k \\ j & j' \end{pmatrix} \rangle_A - (2\beta)^2 \sum_{\substack{k_1 \\ j_1}} \langle V \begin{pmatrix} -k & k_1 \\ j & j_1 \end{pmatrix} \alpha_{\ell(j_1)}^{(k_1)} V \begin{pmatrix} k_1 & k \\ j_1 & j_1 \end{pmatrix} \rangle_A + \dots \right] \\
&\quad \left. \times \beta_{\ell(j')}^{(k)} \right\} \quad (7.1.32)
\end{aligned}$$

By introducing the proper self-energy functions $G_{\ell}(kjj')$, which is given by

$$G_{\ell}\left(\begin{smallmatrix} k \\ jj' \end{smallmatrix}\right) = 2\beta < V\left(\begin{smallmatrix} -k & k \\ j & j' \end{smallmatrix}\right) >_A - (2\beta)^2 \\ \times \sum_{\substack{k_1 \\ j_1}}' < V\left(\begin{smallmatrix} -k & k_1 \\ j & j_1 \end{smallmatrix}\right) \alpha_{\ell}\left(\begin{smallmatrix} k_1 \\ j_1 \end{smallmatrix}\right) V\left(\begin{smallmatrix} -k_1 & k \\ j_1 & j' \end{smallmatrix}\right) >_A + \dots$$

(7.1.33)

where the prime on the sum has the same significance here as in Eq. (5.2.26), we can write $A_{\ell}(kjj')$ as

$$A_{\ell}\left(\begin{smallmatrix} k \\ jj' \end{smallmatrix}\right) = \delta_{jj'} \alpha_{\ell}\left(\begin{smallmatrix} k \\ j \end{smallmatrix}\right) - \beta_{\ell}\left(\begin{smallmatrix} k \\ j \end{smallmatrix}\right) \left\{ G_{\ell}\left(\begin{smallmatrix} k \\ jj' \end{smallmatrix}\right) \right. \\ \left. - \sum_{j_1} G_{\ell}\left(\begin{smallmatrix} k \\ jj_1 \end{smallmatrix}\right) \alpha_{\ell}\left(\begin{smallmatrix} k \\ j_1 \end{smallmatrix}\right) G_{\ell}\left(\begin{smallmatrix} k \\ j_1 j' \end{smallmatrix}\right) + \dots \right\} \beta_{\ell}\left(\begin{smallmatrix} k \\ j_1 \end{smallmatrix}\right)$$

(7.1.34)

Let us denote the function in curly brackets in Eq. (7.1.34) by $C_{\ell}(kjj')$, so that

$$A_{\ell}\left(\begin{smallmatrix} k \\ jj' \end{smallmatrix}\right) = \delta_{jj'} \alpha_{\ell}\left(\begin{smallmatrix} k \\ j \end{smallmatrix}\right) - \beta_{\ell}\left(\begin{smallmatrix} k \\ j \end{smallmatrix}\right) C_{\ell}\left(\begin{smallmatrix} k \\ jj' \end{smallmatrix}\right) \beta_{\ell}\left(\begin{smallmatrix} k' \\ j' \end{smallmatrix}\right)$$

(7.1.35)

From its definition,

$$C_{\ell}\left(\begin{smallmatrix} k \\ jj' \end{smallmatrix}\right) = G_{\ell}\left(\begin{smallmatrix} k \\ jj' \end{smallmatrix}\right) - \sum_{j_1} G_{\ell}\left(\begin{smallmatrix} k \\ jj_1 \end{smallmatrix}\right) \alpha_{\ell}\left(\begin{smallmatrix} k \\ j_1 \end{smallmatrix}\right) G_{\ell}\left(\begin{smallmatrix} k \\ j_1 j' \end{smallmatrix}\right) \\ + \sum_{\substack{j_1 \\ j_2}} G_{\ell}\left(\begin{smallmatrix} k \\ jj_1 \end{smallmatrix}\right) \alpha_{\ell}\left(\begin{smallmatrix} k \\ j_1 \end{smallmatrix}\right) G_{\ell}\left(\begin{smallmatrix} k \\ j_1 j_2 \end{smallmatrix}\right) \alpha_{\ell}\left(\begin{smallmatrix} k \\ j_2 \end{smallmatrix}\right) G_{\ell}\left(\begin{smallmatrix} k \\ j_2 j' \end{smallmatrix}\right) - \dots$$

(7.1.36)

we see that $C_{\ell}(kj j')$ is the solution of the $3r \times 3r$ set of equations

$$C_{\ell}(kj j') = G_{\ell}(kj j') - \sum_{j_1} G_{\ell}(kj j_1) \alpha_{\ell}(j_1) C_{\ell}(kj_1 j') \quad (7.1.37)$$

If we are content with a result which is correct to terms linear in p , we find just as in Sec. V-2 that

$$\begin{aligned} G_{\ell}^{(1)}(kj j') &= \frac{2\beta\hbar}{4} p\lambda \sqrt{\omega_j(k)\omega_{j'}(k)} e^*\left(+\left|\begin{smallmatrix} k \\ j \end{smallmatrix}\right.\right) \cdot e\left(+\left|\begin{smallmatrix} k \\ j' \end{smallmatrix}\right.\right) \\ &- \frac{4\beta^2\hbar^2}{4^2} \frac{1}{N} p\lambda^2 \sqrt{\omega_j(k)\omega_{j'}(k)} \sum_{k_1 j_1} e^*\left(+\left|\begin{smallmatrix} k \\ j \end{smallmatrix}\right.\right) \cdot e\left(+\left|\begin{smallmatrix} k_1 \\ j_1 \end{smallmatrix}\right.\right) \omega_{j_1}(k_1) \alpha_{\ell}(j_1) \\ &\times e^*\left(+\left|\begin{smallmatrix} k_1 \\ j_1 \end{smallmatrix}\right.\right) \cdot e\left(+\left|\begin{smallmatrix} k \\ j' \end{smallmatrix}\right.\right) \\ &+ \frac{8\beta^3\hbar^3}{4^3} \frac{1}{N^2} p\lambda^3 \sqrt{\omega_j(k)\omega_{j'}(k)} \sum_{k_1 k_2 j_1 j_2} e^*\left(+\left|\begin{smallmatrix} k \\ j \end{smallmatrix}\right.\right) \cdot e\left(+\left|\begin{smallmatrix} k_1 \\ j_1 \end{smallmatrix}\right.\right) \omega_{j_1}(k_1) \alpha_{\ell}(j_1) \\ &\times e^*\left(+\left|\begin{smallmatrix} k_1 \\ j_1 \end{smallmatrix}\right.\right) \cdot e\left(+\left|\begin{smallmatrix} k_2 \\ j_2 \end{smallmatrix}\right.\right) \omega_{j_2}(k_2) \alpha_{\ell}(j_2) e^*\left(+\left|\begin{smallmatrix} k_2 \\ j_2 \end{smallmatrix}\right.\right) \cdot e\left(+\left|\begin{smallmatrix} k \\ j' \end{smallmatrix}\right.\right) - \dots \\ &= p\lambda \frac{\beta\hbar}{2} \sqrt{\omega_j(k)\omega_{j'}(k)} \sum_{\alpha\beta} e_{\alpha}^*\left(+\left|\begin{smallmatrix} k \\ j \end{smallmatrix}\right.\right) \left\{ \delta_{\alpha\beta} - \frac{\beta\hbar\lambda}{2} K_{\alpha\beta}^{(\ell)} \right. \\ &\left. + \frac{\beta^2\hbar^2\lambda^2}{2^2} \sum_{\gamma} K_{\alpha\gamma}^{(\ell)} K_{\gamma\beta}^{(\ell)} - \dots \right\} e_{\beta}\left(+\left|\begin{smallmatrix} k \\ j' \end{smallmatrix}\right.\right) \end{aligned} \quad (7.1.38)$$

where we have introduced the matrix $K^{(\ell)}$ whose elements are given by

$$K_{\alpha\beta}^{(\ell)} = \frac{1}{N} \sum_{kj} e_{\alpha}\left(+\left|\begin{smallmatrix} k \\ j \end{smallmatrix}\right.\right) \omega_j(k) \alpha_{\ell}(j) e_{\beta}^*\left(+\left|\begin{smallmatrix} k \\ j \end{smallmatrix}\right.\right) \quad (7.1.39)$$

In the case of particular interest to us, a cubic crystal every atom of which is at a center of inversion symmetry, $K_{\alpha\beta}^{(\ell)}$ becomes isotropic,

$$K_{\alpha\beta}^{(\ell)} = \delta_{\alpha\beta} K_{\ell} = \frac{\delta_{\alpha\beta}}{3N} \sum_{kj} e\left(+\left|\begin{smallmatrix} k \\ j \end{smallmatrix}\right.\right) e^*\left(+\left|\begin{smallmatrix} k \\ j \end{smallmatrix}\right.\right) \times w_j(k) \alpha_{\ell}\left(\begin{smallmatrix} k \\ j \end{smallmatrix}\right) \quad (7.1.40)$$

and we obtain

$$G_{\ell}^{(1)}\left(\begin{smallmatrix} k \\ jj' \end{smallmatrix}\right) = p \frac{\beta \hbar \lambda}{2} \sqrt{w_j(k) w_{j'}(k)} \times \frac{e^*\left(+\left|\begin{smallmatrix} k \\ j \end{smallmatrix}\right.\right) \cdot e\left(+\left|\begin{smallmatrix} k \\ j' \end{smallmatrix}\right.\right)}{1 + \frac{\beta \hbar \lambda}{2} K_{\ell}} \quad (7.1.41)$$

In another special case, a cubic Bravais lattice, we achieve an even greater simplification. By suppressing the basis index + and by using the orthonormality of the eigenvectors, we find

$$G_{\ell}^{(1)}\left(\begin{smallmatrix} k \\ jj' \end{smallmatrix}\right) = \delta_{jj'} p \frac{\lambda \beta \hbar}{2} \frac{w_j(k)}{1 + \frac{\lambda \beta \hbar}{2} K_{\ell}} = \delta_{jj'} G_{\ell}\left(\begin{smallmatrix} k \\ j \end{smallmatrix}\right) \quad (7.1.42)$$

where K_{ℓ} is now given by

$$\begin{aligned}
 K_{\ell} &= \frac{1}{3N} \sum_{\mathbf{kj}} \omega_{\mathbf{j}}(\mathbf{k}) \alpha_{\ell}(\mathbf{k}) \\
 &= \frac{2}{\beta \hbar} \frac{1}{3N} \sum_{\mathbf{kj}} \frac{\omega_{\mathbf{j}}^2(\mathbf{k})}{\omega_{\mathbf{j}}^2(\mathbf{k}) + \omega_{\ell}^2}
 \end{aligned}
 \tag{7.1.43}$$

In this case we can solve for $A_{\ell}(\mathbf{kj}j')$ with the result that

$$A_{\ell}(\mathbf{k}) = \delta_{jj'} \left\{ \alpha_{\ell}(\mathbf{k}) - \beta_{\ell}^2(\mathbf{k}) G_{\ell}(\mathbf{k}) \frac{1}{1 + \alpha_{\ell}(\mathbf{k}) G_{\ell}(\mathbf{k})} \right\}
 \tag{7.1.44}$$

By substituting into this expression the explicit results for $\alpha_{\ell}(\mathbf{kj})$, $\beta_{\ell}(\mathbf{kj})$, and $G_{\ell}(\mathbf{kj})$, we obtain finally

$$\begin{aligned}
 A_{\ell}(\mathbf{k}) &= \delta_{jj'} \left\{ \frac{2\omega_{\mathbf{j}}(\mathbf{k})}{\beta \hbar} \frac{1}{\omega_{\mathbf{j}}^2(\mathbf{k}) + \omega_{\ell}^2} + p\lambda \frac{2\omega_{\ell}^2}{\beta \hbar} \frac{1}{\omega_{\mathbf{j}}^2(\mathbf{k}) + \omega_{\ell}^2} \right. \\
 &\quad \times \left. \frac{\omega_{\mathbf{j}}(\mathbf{k})}{1 + \frac{\lambda \beta \hbar}{2} K_{\ell}} \frac{1}{\omega_{\mathbf{j}}^2(\mathbf{k}) + \omega_{\ell}^2 + p\lambda \frac{\omega_{\mathbf{j}}^2(\mathbf{k})}{1 + \frac{\lambda \beta \hbar}{2} K_{\ell}}} \right\}
 \end{aligned}
 \tag{7.1.45}$$

For future reference we point out that since the correlation function $G_{\alpha\beta}(\ell, \ell'; u)$ is given by

$$\begin{aligned}
 G_{\alpha\beta}(\ell, \ell'; u) &= \langle T u_{\alpha}(\ell; u) u_{\beta}(\ell'; 0) \rangle \\
 &= \frac{\hbar}{2N \sqrt{M_{\alpha} M_{\beta}}} \sum_{\substack{\mathbf{k}, \mathbf{k}' \\ \mathbf{j}, \mathbf{j}'}} \frac{e_{\alpha}(\mathbf{k} | \mathbf{j}) e_{\beta}^{*}(\mathbf{k}' | \mathbf{j}')}{(w_{\mathbf{j}}(\mathbf{k}) w_{\mathbf{j}'}(\mathbf{k}'))^{\frac{1}{2}}} e^{2\pi i (\mathbf{k} \cdot \mathbf{x}(\ell) - \mathbf{k}' \cdot \mathbf{x}(\ell'))} \\
 &\quad \times \sum_{n=-\infty}^{\infty} A_n(\mathbf{k}, \mathbf{k}' | \mathbf{j}, \mathbf{j}') e^{\frac{2\pi i n u}{\beta}}
 \end{aligned}
 \tag{7.1.46}$$

if we pass to the limit $u \rightarrow 0+$ we obtain the equal time correlation function

$$\begin{aligned}
 \langle u_{\alpha}(\ell; 0) u_{\beta}(\ell'; 0) \rangle &= \frac{\hbar}{2N \sqrt{M_{\alpha} M_{\beta}}} \sum_{\substack{\mathbf{k}, \mathbf{k}' \\ \mathbf{j}, \mathbf{j}'}} \frac{e_{\alpha}(\mathbf{k} | \mathbf{j}) e_{\beta}^{*}(\mathbf{k}' | \mathbf{j}')}{(w_{\mathbf{j}}(\mathbf{k}) w_{\mathbf{j}'}(\mathbf{k}'))^{\frac{1}{2}}} \\
 &\quad \times e^{2\pi i (\mathbf{k} \cdot \mathbf{x}(\ell) - \mathbf{k}' \cdot \mathbf{x}(\ell'))} \sum_{n=-\infty}^{\infty} A_n(\mathbf{k}, \mathbf{k}' | \mathbf{j}, \mathbf{j}')
 \end{aligned}
 \tag{7.1.47}$$

We remark that there is no difficulty in passing to the limit $u \rightarrow 0+$. Unlike the Fourier series (7.1.23) that defines a function of u discontinuous at $u = 0$, the Fourier series (7.1.46) defines a function which is continuous at $u = 0$, because the equal time commutator of $u_{\alpha}(\ell, \ell')$ and $u_{\beta}(\ell', \ell')$ vanishes. An important special case of this result is the

case when the atom (l, κ) and the atom (l', κ') are one and the same. In this case we find

$$\begin{aligned} \langle u_{\alpha} \left(\begin{smallmatrix} l \\ \kappa \end{smallmatrix}; 0 \right) u_{\beta} \left(\begin{smallmatrix} l' \\ \kappa' \end{smallmatrix}; 0 \right) \rangle &= \frac{\hbar}{2NM} \sum_{\substack{\mathbf{k}, \mathbf{k}' \\ \mathbf{j}, \mathbf{j}'}} \frac{e_{\alpha} \left(\begin{smallmatrix} \kappa, \mathbf{k} \\ \mathbf{j} \end{smallmatrix} \right) e_{\beta}^* \left(\begin{smallmatrix} \kappa', \mathbf{k}' \\ \mathbf{j}' \end{smallmatrix} \right)}{\left(w_{\mathbf{j}}(\mathbf{k}) w_{\mathbf{j}'}(\mathbf{k}') \right)^{\frac{1}{2}}} e^{2\pi i (\mathbf{k} - \mathbf{k}') \cdot \mathbf{x}(l)} \\ &\times \sum_{n=-\infty}^{\infty} A_n \left(\begin{smallmatrix} \mathbf{k}, \mathbf{k}' \\ \mathbf{j}, \mathbf{j}' \end{smallmatrix} \right) \end{aligned} \quad (7.1.48)$$

2. The Phonon Propagator for a Bravais Crystal Containing a Single Mass Defect

We can take over most of the results of Sec. VII-1 to obtain the phonon propagator for a crystal containing a single mass defect. This result will be required in Secs. IX and X. For simplicity we restrict ourselves to the case of a Bravais crystal.

If we assume that the atom corresponding to $l = 0$ has its mass changed from M to $M' = (1 - \epsilon) M$, the perturbation Hamiltonian is

$$\begin{aligned} V &= \sum_{\alpha} \left\{ \frac{p_{\alpha}^2(0)}{2M'} - \frac{p_{\alpha}^2(0)}{2M} \right\} \\ &= \sum_{\substack{\mathbf{k}_1, \mathbf{k}_2 \\ \mathbf{j}_1, \mathbf{j}_2}} U \left(\begin{smallmatrix} \mathbf{k}_1, \mathbf{k}_2 \\ \mathbf{j}_1, \mathbf{j}_2 \end{smallmatrix} \right) \left[a_{-\mathbf{k}_1, \mathbf{j}_1}^+ a_{\mathbf{k}_2, \mathbf{j}_2} + a_{\mathbf{k}_1, \mathbf{j}_1} a_{-\mathbf{k}_2, \mathbf{j}_2}^+ \right. \\ &\quad \left. - a_{-\mathbf{k}_1, \mathbf{j}_1}^+ a_{-\mathbf{k}_2, \mathbf{j}_2}^+ - a_{\mathbf{k}_1, \mathbf{j}_1} a_{\mathbf{k}_2, \mathbf{j}_2} \right] \end{aligned}$$

(7.2.1)

where

$$U \begin{pmatrix} k_1 k_2 \\ j_1 j_2 \end{pmatrix} = \frac{\lambda \hbar}{4N} \left(\omega_{j_1}(k_1) \omega_{j_2}(k_2) \right)^{\frac{1}{2}} e \begin{pmatrix} k_1 \\ j_1 \end{pmatrix} \cdot e \begin{pmatrix} k_2 \\ j_2 \end{pmatrix} \quad (7.2.2)$$

with

$$\lambda = \frac{\epsilon}{1 - \epsilon} \quad (7.2.3)$$

The desired correlation function becomes

$$\begin{aligned} G_{\alpha\beta}(\ell\ell'; u) &= \langle T u_{\alpha}(\ell; u) u_{\beta}(\ell'; 0) \rangle \\ &= \frac{\hbar}{2NM} \sum_{\substack{k k' \\ j j'}} \frac{e_{\alpha} \begin{pmatrix} k \\ j \end{pmatrix} e_{\beta} \begin{pmatrix} k' \\ j' \end{pmatrix}}{\left(\omega_j(k) \omega_{j'}(k') \right)^{\frac{1}{2}}} e^{2\pi i k \cdot x(\ell) - 2\pi i k' \cdot x(\ell')} \\ &\quad \times F \begin{pmatrix} k k' \\ j j' ; u \end{pmatrix} \end{aligned} \quad (7.2.4)$$

where as before

$$F \begin{pmatrix} k k' \\ j j' ; u \end{pmatrix} = \langle T \left(a_{-k,j}^+ (u) + a_{k,j} (u) \right) \left(a_{-k',j'} (0) + a_{k',j'}^+ (0) \right) \rangle \quad (7.2.5)$$

By comparing these expressions with the corresponding ones in Sec. VII-1, we see that to obtain $A_{\ell}(k_j; k'_{j'})$ in the present case we need only to replace $V(k_j; k'_{j'})$ in Eq. (7.1.31) by $U(k_j; k'_{j'})$. We thus obtain

$$A_{\alpha} \left(\begin{matrix} k & k' \\ j & j' \end{matrix} \right) = \Delta(k-k') \delta_{jj'} \alpha_{\alpha} \left(\begin{matrix} k \\ j \end{matrix} \right) - \beta_{\alpha} \left(\begin{matrix} k \\ j \end{matrix} \right) \left\{ 2\beta U \left(\begin{matrix} k & k' \\ j & j' \end{matrix} \right) - \sum_{\substack{k_1 \\ j_1}} 2\beta U \left(\begin{matrix} k & k_1 \\ j & j_1 \end{matrix} \right) \alpha_{\alpha} \left(\begin{matrix} k_1 \\ j_1 \end{matrix} \right) 2\beta U \left(\begin{matrix} k_1 & k' \\ j_1 & j' \end{matrix} \right) + \dots \right\} \beta_{\alpha} \left(\begin{matrix} k' \\ j' \end{matrix} \right) \quad (7.2.6)$$

In writing this result we have used the fact that $U(k_1 j_1; k_2 j_2)$ is an even function of k_1 and of k_2 .

At least in the case of cubic crystals this result can be summed in closed form. In general, we can express $2\beta U(k_1 j_1; k_2 j_2)$ in the form

$$2\beta U \left(\begin{matrix} k_1 & k_2 \\ j_1 & j_2 \end{matrix} \right) = \sum_{\alpha} f_{\alpha} \left(\begin{matrix} k_1 \\ j_1 \end{matrix} \right) f_{\alpha} \left(\begin{matrix} k_2 \\ j_2 \end{matrix} \right) \quad (7.2.7)$$

where

$$f_{\alpha} \left(\begin{matrix} k \\ j \end{matrix} \right) = \left(\frac{\lambda \beta \hbar}{2N} \right)^{\frac{1}{2}} \sqrt{w_j(k)} e_{\alpha} \left(\begin{matrix} k \\ j \end{matrix} \right) \quad (7.2.8)$$

If we substitute this expression into Eq. (7.2.6) we obtain

$$A_{\alpha} \left(\begin{matrix} k & k' \\ j & j' \end{matrix} \right) = \Delta(k-k') \delta_{jj'} \alpha_{\alpha} \left(\begin{matrix} k \\ j \end{matrix} \right) - \beta_{\alpha} \left(\begin{matrix} k \\ j \end{matrix} \right) \beta_{\alpha} \left(\begin{matrix} k' \\ j' \end{matrix} \right) \sum_{\alpha\beta} f_{\alpha} \left(\begin{matrix} k \\ j \end{matrix} \right) \times \left\{ \delta_{\alpha\beta} - \sum_{\substack{k_1 \\ j_1}} f_{\alpha} \left(\begin{matrix} k \\ j_1 \end{matrix} \right) \alpha_{\alpha} \left(\begin{matrix} k \\ j_1 \end{matrix} \right) f_{\beta} \left(\begin{matrix} k \\ j_1 \end{matrix} \right) \right.$$

$$+ \sum_{\gamma} \sum_{\substack{k_1 k_2 \\ j_1 j_2}} f_{\alpha} \left(\begin{smallmatrix} k_1 \\ j_1 \end{smallmatrix} \right) \alpha_{\ell} \left(\begin{smallmatrix} k_1 \\ j_1 \end{smallmatrix} \right) f_{\gamma} \left(\begin{smallmatrix} k_1 \\ j_1 \end{smallmatrix} \right) f_{\gamma} \left(\begin{smallmatrix} k_2 \\ j_2 \end{smallmatrix} \right) \alpha_{\ell} \left(\begin{smallmatrix} k_2 \\ j_2 \end{smallmatrix} \right) f_{\beta} \left(\begin{smallmatrix} k_2 \\ j_2 \end{smallmatrix} \right) - \dots \left\} f_{\beta} \left(\begin{smallmatrix} k' \\ j' \end{smallmatrix} \right)$$

(7.2.9)

We now introduce a 3×3 matrix G whose elements are

$$G_{\alpha\beta} = \sum_{\substack{k \\ j}} f_{\alpha} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) \alpha_{\ell} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) f_{\beta} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right)$$

(7.2.10)

In terms of this matrix we can rewrite Eq. (7.2.9) as

$$A_{\ell} \left(\begin{smallmatrix} k k' \\ j j' \end{smallmatrix} \right) = \Delta(k-k') \delta_{jj'} \alpha_{\ell} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) - \beta_{\ell} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) \beta_{\ell} \left(\begin{smallmatrix} k' \\ j' \end{smallmatrix} \right) \sum_{\alpha\beta} f_{\alpha} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right)$$

$$\times \left\{ I - G + G^2 - \dots \right\}_{\alpha\beta} f_{\beta} \left(\begin{smallmatrix} k' \\ j' \end{smallmatrix} \right)$$

$$= \Delta(k-k') \delta_{jj'} \alpha_{\ell} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) - \beta_{\ell} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) \beta_{\ell} \left(\begin{smallmatrix} k' \\ j' \end{smallmatrix} \right)$$

$$\times \sum_{\alpha\beta} f_{\alpha} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) \left[(I + G)^{-1} \right]_{\alpha\beta} f_{\beta} \left(\begin{smallmatrix} k' \\ j' \end{smallmatrix} \right)$$

(7.2.11)

The element $G_{\alpha\beta}$ written out explicitly is

$$G_{\alpha\beta} = \frac{\lambda}{N} \sum_{kj} \frac{e_{\alpha} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right) \omega_j^2(k) e_{\beta} \left(\begin{smallmatrix} k \\ j \end{smallmatrix} \right)}{\omega_{\ell}^2 + \omega_j^2(k)}$$

$$= \lambda \delta_{\alpha\beta} - \lambda \omega_{\ell}^2 \frac{1}{N} \sum_{kj} \frac{e_{\alpha}^{(k)}(j) e_{\beta}^{(k)}(j)}{\omega_{\ell}^2 + \omega_j^2(k)}$$

(7. 2. 12)

In the special case of a cubic crystal $G_{\alpha\beta}$ is isotropic:

$$G_{\alpha\beta} = \lambda \delta_{\alpha\beta} \left\{ 1 - \frac{\omega_{\ell}^2}{3N} \sum_{kj} \frac{1}{\omega_{\ell}^2 + \omega_j^2(k)} \right\} = \lambda \delta_{\alpha\beta} G_{\ell}$$

(7. 2. 13)

In this case the expression for $A_{\ell}(kj; k'j')$ simplifies to

$$\begin{aligned} A_{\ell} \left(\begin{matrix} k & k' \\ j & j' \end{matrix} \right) &= \Delta(k-k') \delta_{jj'} \alpha_{\ell} \left(\begin{matrix} k \\ j \end{matrix} \right) - \frac{\lambda \beta \hbar}{2N} \sqrt{\omega_j(k) \omega_{j'}(k')} \\ &\times \beta_{\ell} \left(\begin{matrix} k \\ j \end{matrix} \right) \frac{e \left(\begin{matrix} k \\ j \end{matrix} \right) \cdot e \left(\begin{matrix} k' \\ j' \end{matrix} \right)}{1 + \lambda G_{\ell}} \beta_{\ell} \left(\begin{matrix} k' \\ j' \end{matrix} \right) \\ &= \Delta(k-k') \delta_{jj'} \frac{2\omega_j(k)}{\beta \hbar} \frac{1}{\omega_j^2(k) + \omega_{\ell}^2} \\ &+ \frac{\lambda \beta \hbar}{2N} \sqrt{\omega_j(k) \omega_{j'}(k')} \frac{2\omega_{\ell}}{\beta \hbar} \frac{1}{\omega_j^2(k) + \omega_{\ell}^2} \frac{e \left(\begin{matrix} k \\ j \end{matrix} \right) \cdot e \left(\begin{matrix} k' \\ j' \end{matrix} \right)}{1 + \lambda G_{\ell}} \\ &\times \frac{2\omega_{\ell}}{\beta \hbar} \frac{1}{\omega_{j'}^2(k') + \omega_{\ell}^2} \end{aligned}$$

(7. 2. 14)

VIII. OPTICAL ABSORPTION IN ISOTOPICALLY DISORDERED CRYSTALS

1. Formula for the Optical Absorption Coefficient

In this section we apply the results of Sec. VII to the calculation of the optical absorption in the infrared by mixed alkali-halide crystals or by isotopically disordered alkali-halide crystals.

For the purpose of simplifying the calculation we assume the following simple model. To a pure NaCl lattice (say) we add substitutionally some KCl (say). The sublattice of chloride ions will be assumed to remain perfectly ordered and the sublattice of the positive ions is now randomly disordered. Alternatively we could consider infrared absorption in a LiF crystal, for example, in which some of the ${}^7\text{Li}^+$ ions have been replaced randomly by ${}^6\text{Li}^+$ ions.

In deriving the expression for the optical absorption coefficient which forms the basis for the work described in this section, we follow a procedure similar to that employed by Kubo in his treatment of the magnetic susceptibility.⁵⁴ Our starting point is the quantum statistical mechanical result that the ensemble average of the expectation value of an operator B is given by

$$\langle B \rangle = \text{Tr } \rho B \quad (8.1.1)$$

where ρ is the density matrix for the ensemble. In the absence of an external time-dependent perturbation, ρ satisfies the equation of motion

$$\frac{\partial \rho}{\partial t} = \frac{1}{i\hbar} [H, \rho] \quad (8.1.2)$$

where H is the Hamiltonian of the system of interest, in our case an isotopically disordered crystal. We now assume that at time $t = -\infty$ an external electromagnetic field $E(t)$ is turned on. The interaction energy of the crystal with the external field is given by

$$H_I(t) = -M \cdot E(t) \quad (8.1.3)$$

where M is the crystal's dipole moment. The equation of motion of the perturbed density matrix $\rho + \Delta\rho$ is

$$\begin{aligned}\frac{\partial}{\partial t}(\rho + \Delta\rho) &= \frac{1}{i\hbar} [H - M \cdot E(t), \rho + \Delta\rho] \\ &\approx \frac{1}{i\hbar} \{ [H, \rho] + [H, \Delta\rho] - [M \cdot E(t), \rho] \}\end{aligned}\quad (8.1.4)$$

where, since we shall only be interested in a linear relation between the crystal's dipole moment and the external field, we have neglected the commutator $[M \cdot E(t), \Delta\rho]$ as small of the second order in $E(t)$. The equation of motion of $\Delta\rho$ is thus given by

$$\frac{\partial \Delta\rho}{\partial t} = \frac{1}{i\hbar} \{ [H, \Delta\rho] - [M \cdot E(t), \rho] \} \quad (8.1.5)$$

Kubo has solved this equation subject to the initial condition $\Delta\rho(-\infty) = 0$, with the result that

$$\Delta\rho = \frac{i}{\hbar} \sum_{\nu} \int_{-\infty}^t e^{-i(t-t')\frac{H}{\hbar}} [M_{\nu}, \rho] e^{i(t-t')\frac{H}{\hbar}} E_{\nu}(t') dt' \quad (8.1.6)$$

The ensemble average of the expectation value of the μ component of the crystal's dipole moment is given by

$$\begin{aligned}\langle M_{\mu} \rangle &= \text{Tr} (\rho + \Delta\rho) M_{\mu} \\ &= \text{Tr} \Delta\rho M_{\mu} \\ &= \text{Tr} \frac{i}{\hbar} \sum_{\nu} \int_{-\infty}^t e^{-i(t-t')\frac{H}{\hbar}} [M_{\nu}, \rho] e^{i(t-t')\frac{H}{\hbar}} E_{\nu}(t') M_{\mu} dt'\end{aligned}\quad (8.1.7)$$

since we assume that the crystal has no permanent moment in the absence of the external field. If we introduce the notation

$$M(t) = e^{i\frac{t}{\hbar}H} M e^{-i\frac{t}{\hbar}H} \quad (8.1.8)$$

and use the cyclic theorem for traces,

$$\text{Tr } ABC = \text{Tr } BCA, \text{ etc.}$$

we can rewrite $\langle M_\mu \rangle$ as

$$\langle M_\mu \rangle = \text{Tr} \frac{i}{\hbar} \sum_v \int_0^\infty [M_{v,\rho}] M_\mu(\tau) E_v(t-\tau) d\tau \quad (8.1.9)$$

We describe the external field by

$$E(t) = e^{\epsilon t} E_0 e^{i\omega t} \quad (8.1.10)$$

where the factor $e^{\epsilon t}$ is introduced for the purpose of turning on the field adiabatically for $t < 0$; at the end of our calculations, we pass to the limit $\epsilon \rightarrow 0+$. We finally have

$$\begin{aligned} \langle M_\mu \rangle = & \frac{i}{\hbar} \text{Tr} \sum_v E_{0v} e^{\epsilon t + i\omega t} \\ & \times \int_0^\infty [M_{v,\rho}] M_\mu(\tau) e^{-\epsilon\tau - i\omega\tau} d\tau \end{aligned} \quad (8.1.11)$$

so that the $\mu\nu$ element of the dielectric susceptibility tensor is given by

$$\chi_{\mu\nu}(\omega) = \frac{i}{\hbar V} \lim_{\epsilon \rightarrow 0+} \text{Tr} \int_0^{\infty} e^{-\epsilon t - i\omega t} [M_{\nu}, \rho] M_{\mu}(t) dt \quad (8.1.12)$$

where V is the volume of the crystal.

We assume that the unperturbed crystal is in a canonical distribution so that

$$\rho = \frac{e^{-\beta H}}{Z} \quad (8.1.13)$$

where Z is the crystal's partition function. We can use the cyclic theorem for traces to rewrite Eq. (8.1.12) as

$$\chi_{\mu\nu}(\omega) = \frac{i}{\hbar V} \lim_{\epsilon \rightarrow 0+} \text{Tr} \int_0^{\infty} e^{-\epsilon t - i\omega t} \times \rho [M_{\mu}(t), M_{\nu}] dt \quad (8.1.14)$$

We shall be concerned only with cubic crystals in what follows. In this case the susceptibility tensor becomes isotropic and we need to consider only the xx element

$$\chi_{xx}(\omega) = \frac{i}{\hbar V} \lim_{\epsilon \rightarrow 0+} \text{Tr} \int_0^{\infty} e^{-\epsilon t - i\omega t} \times \rho [M_x(t), M_x] dt \quad (8.1.15)$$

Since the operators $M_x(t)$ and M_x are Hermitian, we find readily that

$$\begin{aligned} \langle [M_x(t), M_x] \rangle &= \text{Tr } \rho [M_x(t), M_x] \\ &= 2i \text{Im} \langle M_x(t) M_x(0) \rangle \end{aligned} \quad (8.1.16)$$

and Eq. (8.1.15) becomes

$$\begin{aligned} \chi_{xx}(\omega) &= \frac{2}{\hbar V} \lim_{\epsilon \rightarrow 0+} \int_0^{\infty} e^{-\epsilon t - i\omega t} \\ &\quad \times \text{Im} \langle M_x(t) M_x(0) \rangle dt \end{aligned} \quad (8.1.17)$$

If the susceptibility tensor is written

$$\chi_{xx}(\omega) = \chi_{xx}^{(1)}(\omega) - i \chi_{xx}^{(2)}(\omega) \quad (8.1.18)$$

then the xx component of the absorption coefficient is given by

$$\alpha_{xx}(\omega) = 4\pi \frac{\omega}{c} \chi_{xx}^{(2)}(\omega) \quad (8.1.19)$$

where c is the speed of light. By combining Eqs. (8.1.17) and (8.1.19), we can write the absorption coefficient as

$$\begin{aligned} \alpha_{xx}(\omega) &= -\frac{4\pi\omega}{c} \frac{2}{\hbar V} \lim_{\epsilon \rightarrow 0+} \int_0^{\infty} e^{-\epsilon t} \sin \omega t \\ &\quad \times \text{Im} \langle M_x(t) M_x(0) \rangle dt \end{aligned} \quad (8.1.20)$$

To emphasize the isotropic nature of the absorption coefficient we can express this result alternatively as

$$\begin{aligned}\alpha(\omega) &= \frac{1}{3} \sum_{\mathbf{x}} \alpha_{\mathbf{x}\mathbf{x}}(\omega) \\ &= -\frac{8\pi\omega}{3c} \frac{1}{\hbar V} \operatorname{Im} \lim_{\epsilon \rightarrow 0+} \int_0^{\infty} e^{-\epsilon t} \sin \omega t \langle M(t) \cdot M(0) \rangle dt\end{aligned}$$

(8.1.21)

2. The Absorption Coefficient and the Propagator Formalism

Let us denote the correlation function $\langle M(t) \cdot M(0) \rangle$ by $F(t)$. If we expand this function in terms of the eigenstates $|m\rangle$ of the crystal's Hamiltonian H , we can write the Fourier sine transform of $F(t)$ as

$$\begin{aligned}F(\omega) &= \int_0^{\infty} dt e^{-\epsilon t} \sin \omega t F(t) \\ &= \frac{1}{2Z} \sum_{mn} e^{-\beta E_m} | \langle n | M(0) | m \rangle |^2 \\ &\quad \times \left\{ \frac{1}{\omega + \frac{1}{\hbar} (E_m - E_n) + i\epsilon} + \frac{1}{\omega - \frac{1}{\hbar} (E_m - E_n) - i\epsilon} \right\}\end{aligned}$$

(8.2.1)

If we now pass to the limit $\epsilon \rightarrow 0+$ we can express the imaginary part of $F(\omega)$ as

$$\operatorname{Im} \lim_{\epsilon \rightarrow 0+} F(\omega) = -\frac{\pi}{2} (e^{\beta \hbar \omega} - 1) \rho(\omega) \quad (8.2.2)$$

where

$$\rho(\omega) = \frac{1}{Z} \sum_{mn} e^{-\beta E_m} |\langle n | M(0) | m \rangle|^2 \delta\left(\omega - \frac{1}{\hbar} (E_m - E_n)\right) \quad (8.2.3)$$

By combining Eqs. (8.1.21) and (8.2.2), we find that the absorption coefficient can be expressed as

$$\alpha(\omega) = \frac{4\pi^2}{3c} \frac{\omega}{\hbar V} (e^{\beta \hbar \omega} - 1) \rho(\omega) \quad (8.2.4)$$

The absorption coefficient has a simple formal appearance in terms of the spectral function $\rho(\omega)$. To calculate $\rho(\omega)$ we proceed as follows. We introduce a function $f(u)$ which is defined by

$$f(u) = \langle T e^{uH} M(0) e^{-uH} \cdot M(0) \rangle \quad (8.2.5)$$

for real u . If we express $f(u)$ in terms of the eigenstates of H we find that it has the property

$$f(u + \beta) = f(u) \quad -\beta < u < 0 \quad (8.2.6)$$

This means that we only have to determine $f(u)$ for u in the interval $-\beta < u < 0$ to know its value outside this interval.

It is convenient to expand $f(u)$ in a Fourier series with a period β :

$$f(u) = \sum_l A_l e^{\frac{2\pi i l u}{\beta}} \quad (8.2.7)$$

The Fourier coefficients $\{A_l\}$ are given by

$$a_l = \frac{1}{\beta} \int_0^\beta e^{\frac{-2\pi i l u}{\beta}} f(u) du \quad (8.2.8)$$

$$a_l = \frac{1}{\beta \hbar Z} \sum_{mn} e^{-\beta E_m} \left(e^{\beta(E_m - E_n)} - 1 \right) | \langle n | M(0) | m \rangle |^2 \\ \times \frac{1}{\frac{1}{\hbar} (E_m - E_n) - \frac{2\pi i l}{\beta \hbar}} \quad (8.2.9)$$

We now define a function $a(\nu)$ of a continuous variable such that

$$a(\nu) = a_l \quad \text{when} \quad \nu = \nu_l = \frac{2\pi i l}{\beta \hbar} \quad (8.2.10)$$

To specify $a(\nu)$ completely, we require that it have a branch cut along the real axis, be analytic everywhere else in the complex ν plane, and go to zero as ν approaches infinity along any straight line in the upper or lower half plane.^{54a} We see from Eq. (8.2.9) that we can obtain $a(\nu)$ simply by replacing the discrete variable $2\pi i l / \beta \hbar$ by ν .

We then find

$$\lim_{\delta \rightarrow 0+} \frac{a(\nu + i\delta) - a(\nu - i\delta)}{2\pi i} = \frac{e^{\beta \hbar \nu} - 1}{\beta \hbar Z} \sum_{mn} e^{-\beta E_m} \\ \times | \langle n | M(0) | m \rangle |^2 \delta \left(\frac{1}{\hbar} (E_m - E_n) - \nu \right) \\ = \frac{1}{\beta \hbar} (e^{\beta \hbar \nu} - 1) \rho(\nu)$$

Thus we have the result

$$\rho(\nu) = \frac{\beta \hbar}{e^{\beta \hbar \nu} - 1} \lim_{\delta \rightarrow 0+} \frac{a(\nu + i\delta) - a(\nu - i\delta)}{2\pi i} \quad (8.2.11)$$

By combining Eqs. (8.2.4) and (8.2.11), we obtain the desired result:

$$\alpha(\omega) = \frac{4\pi^2 \omega}{3c} \frac{\beta}{V} \lim_{\delta \rightarrow 0+} \frac{a(\omega + i\delta) - a(\omega - i\delta)}{2\pi i} \quad (8.2.12)$$

The particular advantage of this form for the absorption coefficient is that the calculation of the Fourier coefficients $\{a_\nu\}$ and hence the function $a(\nu)$, can be reduced to a calculation of the displacement correlation function, which we carried out in the preceding section.

3. Result for the Absorption Coefficient

The dipole moment of an alkali-halide crystal is given by

$$M = \sum_{l, \kappa} \epsilon_\kappa u_\kappa^{(l)} \quad (8.3.1)$$

where ϵ_κ is the charge on the κ -th kind of ion. We express this result in terms of phonon creation and destruction operators using Eq. (7.1.6a):

$$M = \left(\frac{\hbar}{2N}\right)^{\frac{1}{2}} \sum_{l, \kappa} \frac{\epsilon_\kappa}{\sqrt{M_\kappa}} \sum_{kj} \frac{e^{i(\kappa | j | k)}}{\sqrt{\omega_j(k)}} \left(a_{-kj}^\dagger + a_{kj} \right) e^{2\pi i k \cdot x(l)}$$

$$= \left(\frac{N\hbar}{2} \right)^{\frac{1}{2}} \sum_j \frac{M_j^{(o)}}{\{w_j(o)\}^{\frac{1}{2}}} (a_{oj}^+ + a_{oj}) \quad (8.3.2)$$

where

$$M_j^{(o)} = \sum_{\kappa} \frac{e_{\kappa}}{\sqrt{M_{\kappa}}} e(\kappa | j^o) \quad (8.3.3)$$

It can be shown that $M(j^o)$ vanishes when j refers to any of the acoustic branches.⁵⁵ Consequently the sum over j in Eq. (8.3.2) runs over the optical branches only. A more careful analysis of the optical absorption problem by Wallis and Maradudin⁵⁶ leads to the conclusion that the sum over j in Eq. (8.3.2) must further be restricted to either of the two transverse optical branches. In what follows we shall use this result.

According to Eq. (8.2.12) in order to obtain the absorption coefficient, we must obtain the discontinuity across the real axis of the function $a(\omega)$, which is the function of the continuous variable ω obtained from the Fourier coefficient of the correlation function

$$\begin{aligned} f(u) &= \langle T e^{uH} M(o) e^{-uH} \cdot M(o) \rangle \\ &= \frac{N\hbar}{2w_j(o)} M_j^{(o)} \cdot M_j^{(o)} \langle T e^{uH} (a_{oj}^+ + a_{oj}) e^{-uH} (a_{oj}^+ + a_{oj}) \rangle \end{aligned} \quad (8.3.4)$$

The required Fourier coefficient is given with the aid of Eq. (7.1.35)

$$a_l = \frac{N \hbar}{2w_j(0)} M^2 \begin{pmatrix} 0 \\ j \end{pmatrix} \left[\alpha_l \begin{pmatrix} 0 \\ j \end{pmatrix} - \beta_l \begin{pmatrix} 0 \\ j \end{pmatrix} C_{l(jj)} \begin{pmatrix} 0 \\ j \end{pmatrix} \beta_l \begin{pmatrix} 0 \\ j \end{pmatrix} \right]$$

(8.3.5)

If we replace $\alpha_l \begin{pmatrix} 0 \\ j \end{pmatrix}$ and $\beta_l \begin{pmatrix} 0 \\ j \end{pmatrix}$ by their explicit expressions using Eq. (7.1.28) and replace w_l by v/i , according to Eq. (8.2.10), we find that

$$a(v) = \frac{N \hbar M^2 \begin{pmatrix} 0 \\ j \end{pmatrix}}{2w_j(0)} \left[\frac{2w_j(0)}{\beta \hbar} \frac{1}{w_j^2(0) - v^2} - \frac{4v^2}{\beta^2 \hbar^2} \frac{1}{(w_j^2(0) - v^2)^2} C_{l(jj)} \begin{pmatrix} 0 \\ jj \end{pmatrix} ; i w_l = v \right]$$

(8.3.6)

We can simplify the results beyond this point by considering more carefully the expression for $C_l \begin{pmatrix} 0 \\ j j' \end{pmatrix}$. According to Eq. (7.1.37) it is the solution of

$$C_{l(jj')} \begin{pmatrix} 0 \\ j j' \end{pmatrix} = G_{l(jj)} \begin{pmatrix} 0 \\ j j \end{pmatrix} - \sum_{j_1} G_{l(jj_1)} \alpha_{l(j_1)} \begin{pmatrix} 0 \\ j_1 \end{pmatrix} C_{l(j_1 j')} \begin{pmatrix} 0 \\ j_1 j' \end{pmatrix}$$

(8.3.7)

where $G_l \begin{pmatrix} 0 \\ j j_1 \end{pmatrix}$ is given in the present case by Eq. (7.1.41);

$$G_l^{(1)} \begin{pmatrix} 0 \\ jj_1 \end{pmatrix} = p \frac{\lambda}{2} \beta \hbar \sqrt{w_j(0) w_{j_1}(0)}$$

$$\times \frac{e^* \left(+ \left| \begin{smallmatrix} 0 \\ j \end{smallmatrix} \right\rangle \right) \cdot e \left(+ \left| \begin{smallmatrix} 0 \\ j_1 \end{smallmatrix} \right\rangle \right)}{1 + \frac{\lambda}{3N} \sum_{\substack{k' \\ j'}} \frac{e^* \left(+ \left| \begin{smallmatrix} k' \\ j' \end{smallmatrix} \right\rangle \right) \cdot e^* \left(+ \left| \begin{smallmatrix} k' \\ j' \end{smallmatrix} \right\rangle \right) \omega_{j'}^2(k')}{\omega_{j'}^2(k') + \omega_l^2}} \quad (8.3.8)$$

In Eq. (8.3.8), j refers to either of the two transverse optical modes. However, we see from this equation that j_1 must also refer to an optical mode, because $\omega_{j_1}(0)$ vanishes for acoustic modes. This is a convenient result because, as we shall now show, it means that $G_{\ell}^{(1)}(0 \ j \ j_1)$ is diagonal in j and j_1 .

For $k = 0$ the eigenvector $e \left(+ \left| \begin{smallmatrix} k \\ j \end{smallmatrix} \right\rangle \right)$ is real. This follows from the fact that the elements of the dynamical matrix, Eq. (3.1.12) are real for $k = 0$. The orthonormality condition on the eigenvectors of a diatomic crystal, Eq. (3.1.14a), can be expressed as

$$e^* \left(+ \left| \begin{smallmatrix} k \\ j \end{smallmatrix} \right\rangle \right) \cdot e \left(+ \left| \begin{smallmatrix} k \\ j' \end{smallmatrix} \right\rangle \right) + e^* \left(- \left| \begin{smallmatrix} k \\ j \end{smallmatrix} \right\rangle \right) \cdot e \left(- \left| \begin{smallmatrix} k \\ j' \end{smallmatrix} \right\rangle \right) = \delta_{jj'} \quad (8.3.9)$$

and at $k = 0$ becomes

$$e \left(+ \left| \begin{smallmatrix} 0 \\ j \end{smallmatrix} \right\rangle \right) \cdot e \left(+ \left| \begin{smallmatrix} 0 \\ j' \end{smallmatrix} \right\rangle \right) + e \left(- \left| \begin{smallmatrix} 0 \\ j \end{smallmatrix} \right\rangle \right) \cdot e \left(- \left| \begin{smallmatrix} 0 \\ j' \end{smallmatrix} \right\rangle \right) = \delta_{jj'} \quad (8.3.10)$$

Let us assume that j always refers to an optical branch in what follows. In reference 16 it is shown that for an optical branch

$$\sqrt{M_+} e \left(+ \left| \begin{smallmatrix} 0 \\ j \end{smallmatrix} \right\rangle \right) = -\sqrt{M_-} e \left(- \left| \begin{smallmatrix} 0 \\ j \end{smallmatrix} \right\rangle \right) \quad (8.3.11)$$

If we substitute this result into Eq. (8.3.10) we obtain

$$e \left(+ \left| \begin{smallmatrix} 0 \\ j \end{smallmatrix} \right\rangle \right) \cdot \left[e \left(+ \left| \begin{smallmatrix} 0 \\ j' \end{smallmatrix} \right\rangle \right) - \sqrt{\frac{M_+}{M_-}} e \left(- \left| \begin{smallmatrix} 0 \\ j' \end{smallmatrix} \right\rangle \right) \right] = \delta_{jj'} \quad (8.3.12)$$

In the case that j' also refers to an optical branch we have, combining Eqs. (8.3.11) and (8.3.12),

$$e\left(+\left|\begin{smallmatrix} o \\ j \end{smallmatrix}\right.\right) \cdot e\left(+\left|\begin{smallmatrix} o \\ j' \end{smallmatrix}\right.\right) = \frac{M_-}{M_+ + M_-} \delta_{jj'} \quad (8.3.13)$$

for j and j' both referring to optical branches. We can use this result to write Eq. (8.3.8) as

$$\begin{aligned} G_{\ell}^{(1)}\left(\begin{smallmatrix} o \\ j j_1 \end{smallmatrix}\right) &= \delta_{j j_1} \frac{\epsilon p}{2} \beta \hbar \omega_j(o) \frac{M_-}{M_+ + M_-} \\ &\times \left\{ 1 - \frac{\epsilon \omega_{\ell}^2}{3N} \sum_{\substack{k' \\ j'}} \frac{e\left(+\left|\begin{smallmatrix} k' \\ j' \end{smallmatrix}\right.\right) \cdot e^*\left(+\left|\begin{smallmatrix} k' \\ j' \end{smallmatrix}\right.\right)}{\omega_{j', (k')}^2 + \omega_{\ell}^2} \right\}^{-1} \\ &= \delta_{j j_1} g_{\ell} \left(\begin{smallmatrix} o \\ j \end{smallmatrix}\right) \end{aligned} \quad (8.3.14)$$

Equation (8.3.7) can now be solved readily and we obtain

$$C_{\ell}\left(\begin{smallmatrix} o \\ j j \end{smallmatrix}\right) = \frac{g_{\ell}\left(\begin{smallmatrix} o \\ j \end{smallmatrix}\right)}{1 + \alpha_{\ell}\left(\begin{smallmatrix} o \\ j \end{smallmatrix}\right) g_{\ell}\left(\begin{smallmatrix} o \\ j \end{smallmatrix}\right)} \quad (8.3.15)$$

The function $a(\nu)$ given by Eq. (8.3.6) can thus be written explicitly as

$$a(\nu) = N k T M^2 \left(\begin{smallmatrix} o \\ j \end{smallmatrix}\right) \left\{ \frac{1}{\omega_j^2(o) - \nu^2} - \frac{\nu^2}{\omega_j^2(o) - \nu^2} \right\}$$

$$\times \epsilon_p \mu \frac{1}{F(\nu^2)} \left\{ \omega_j^2(0) - \nu^2 + \frac{\epsilon_p \mu \omega_j^2(0)}{F(\nu^2)} \right\}^{-1} \quad (8.3.16)$$

where for the sake of convenience we have put

$$\mu = \frac{M_-}{M_+ + M_-} \quad (8.3.17a)$$

$$F(\nu^2) = 1 - \frac{\epsilon \nu^2}{3N} \sum_{\mathbf{k}} \frac{e\left(+\left|\mathbf{k}\right|_j\right) \cdot e^*\left(+\left|\mathbf{k}\right|_j\right)}{\nu^2 - \omega_j^2(\mathbf{k})} \quad (8.3.17b)$$

The absorption coefficient is therefore finally given by

$$\begin{aligned} \alpha(\omega) = & \frac{4\pi^2\omega}{3c} \frac{M^2\left(\begin{smallmatrix} 0 \\ j \end{smallmatrix}\right)}{v_a} \delta\left(\omega_j^2(0) - \omega^2\right) - \frac{4\pi^2\omega}{3c} \frac{M^2\left(\begin{smallmatrix} 0 \\ j \end{smallmatrix}\right)}{v_a} \frac{1}{\pi} \\ & \times \operatorname{Im} \lim_{\delta \rightarrow 0+} \frac{\omega^2 + i\delta}{\omega_j^2(0) - \omega^2 - i\delta} \\ & \times \frac{\epsilon_p \mu F^{-1}(\omega^2 + i\delta)}{\omega_j^2(0) - \omega^2 - i\delta - \epsilon_p \mu \omega_j^2(0) F^{-1}(\omega^2 + i\delta)} \end{aligned} \quad (8.3.18)$$

The first term on the right side of Eq. (8.3.18) is the well-known result for the absorption spectrum of a perfect ionic crystal in the harmonic approximation. It consists of a single δ function peak at the reststrahl frequency $\omega_j(0)$.

The second term on the right side of Eq. (8.3.18) represents the modification in the absorption spectrum as a result of the disorder. We can express this contribution in compact form if we introduce the functions

$$\lim_{\delta \rightarrow 0+} F(\omega^2 + i\delta) = A(\omega^2) + iB(\omega^2) \quad (8.3.19a)$$

$$\begin{aligned} &= 1 - \frac{e\omega^2}{3N} \sum_{\mathbf{k}} \frac{e\left(+\left|\frac{\mathbf{k}}{j}\right|\right) \cdot e^*\left(+\left|\frac{\mathbf{k}}{j}\right|\right)}{\left(\omega^2 - \omega_j^2(\mathbf{k})\right)_p} \\ &+ i\pi \frac{e\omega^2}{3N} \sum_{\mathbf{k}} e\left(+\left|\frac{\mathbf{k}}{j}\right|\right) \cdot e^*\left(+\left|\frac{\mathbf{k}}{j}\right|\right) \delta\left(\omega^2 - \omega_j^2(\mathbf{k})\right) \end{aligned} \quad (8.3.19b)$$

It is also very convenient to introduce the functions

$$\alpha(\omega^2) = \frac{A(\omega^2)}{A^2(\omega^2) + B^2(\omega^2)} \quad (8.3.20a)$$

$$\beta(\omega^2) = \frac{B(\omega^2)}{A^2(\omega^2) + B^2(\omega^2)} \quad (8.3.20b)$$

In terms of these functions the second term on the right side of Eq. (8.3.18) can be expressed

$$\begin{aligned} \alpha^{(1)}(\omega) &= -\frac{4\pi^2\omega}{3c} \frac{M_j^{(0)}}{v_a} \frac{1}{\pi} \\ &\times \operatorname{Im} \left[\frac{\omega^2}{\left(\omega_j^2(0) - \omega^2\right)_p} + i\pi\omega_j^2(0) \delta\left(\omega^2 - \omega_j^2(0)\right) \right] \end{aligned}$$

$$\begin{aligned}
& \times \frac{(\epsilon \mu \alpha - i \epsilon \mu \beta) [\omega_j^2(\omega) - \omega^2 + \epsilon \mu \omega_j^2(\omega) \alpha + i \epsilon \mu \omega_j^2(\omega) \beta]}{[\omega_j^2(\omega) - \omega^2 + \epsilon \mu \omega_j^2(\omega) \alpha]^2 + \epsilon^2 \mu^2 \omega_j^4(\omega) \beta^2} \\
& = \frac{4\pi\omega^3}{3c} \frac{M_j^{(0)}}{v_a} \frac{\epsilon \mu \beta}{[\omega_j^2(\omega) - \omega^2 + \epsilon \mu \omega_j^2(\omega) \alpha]^2 + \epsilon^2 \mu^2 \omega_j^4(\omega) \beta^2} \\
& \quad - \frac{4\pi^2 \omega_j(\omega)}{3c} \frac{M_j^{(0)}}{v_a} \delta(\omega^2 - \omega_j^2(\omega))
\end{aligned}
\tag{8.3.21}$$

We see from this result that the effect of the disorder on the absorption spectrum of a harmonic ionic crystal is to subtract off the δ function peak at the reststrahl frequency and to replace it by a Lorentzian expression peaked about the reststrahl frequency.

The final expression for the absorption coefficient is, therefore,

$$\alpha(\omega) = \frac{4\pi\omega^3}{3c} \frac{M_j^{(0)}}{v_a} \frac{\epsilon \mu \beta}{[\omega_j^2(\omega) - \omega^2 + \epsilon \mu \omega_j^2(\omega) \alpha]^2 + \epsilon^2 \mu^2 \omega_j^4(\omega) \beta^2}
\tag{8.3.22}$$

This result is nonvanishing only in the interval $(0, \omega_L)$, since

$\beta(\omega^2)$ vanishes outside this interval. If $0 < \epsilon < 1$, and ϵ is sufficiently large that a localized vibration mode appears, then we obtain a contribution to the absorption coefficient for $\omega > \omega_L$. This contribu-

tion is also obtained from the second term on the right side of Eq.

(8.3.17b). In this case $\beta(\omega^2) = 0$, and we obtain

$$\alpha(\omega) = \frac{4\pi^2 \omega^3}{3c} \frac{M_j^{(0)}}{v_a} \frac{1}{\omega^2 - \omega_j^2(0)} \times \frac{\epsilon \mu}{A(\omega^2)} \delta \left(\omega_j^2(0) - \omega^2 + \epsilon \mu \frac{\omega_j^2(0)}{A(\omega^2)} \right) \quad (8.2.23)$$

The value of ω_r^2 for which the argument of the δ function vanishes, ω_r^2 , can be obtained approximately in the following way. The frequency of the localized mode resulting from an isolated impurity, ω_0^2 , is given by the solution of

$$A(\omega^2) = 0 \quad \omega > \omega_L \quad (8.2.24)$$

The equation for ω_r^2 can thus be written approximately as

$$\omega_r^2 - \omega_j^2(0) = \epsilon \mu \omega_j^2(0) \frac{1}{(\omega_r^2 - \omega_j^2(0)) A'(\omega_0^2)} \quad (8.2.25)$$

and has the approximate solution

$$\omega_r^2 \approx \omega_0^2 + \epsilon \mu \omega_j^2(0) \frac{1}{A'(\omega_0^2) (\omega_0^2 - \omega_j^2(0))} \quad (8.2.26)$$

The final expression for $\alpha(\omega)$ in this case can be written

$$\alpha(\omega) = \frac{4\pi^2\omega_r^3}{3c} \frac{M_j^{(0)}}{v_a} \frac{1}{\omega_j^2(\omega)} \times \frac{\delta(\omega^2 - \omega_r^2)}{\left| 1 + \frac{\epsilon p \mu \omega_j^2(\omega) A'(\omega_r^2)}{A^2(\omega_r^2)} \right|} \quad \begin{array}{l} 0 < \epsilon < 1 \\ \omega > \omega_L \end{array}$$

(8. 3. 27)

The peak in the absorption spectrum arising from a light substitutional impurity in an alkali-halide crystal has been observed experimentally by Schaefer.⁵⁷ In his work he replaced halogen ions by H^- ions. Such a defect is called a U-center. Because of the smallness of the impurity ion mass compared with the mass of the ion it replaces, e. g.,

$M'/M \cong 1/35$ in the case of a chloride ion, the impurity mode frequency lies considerably above the reststrahl frequency, and the peak at this frequency shows up distinctly in the absorption spectrum. The concentration of U-centers in Schaefer's work was so low that the difference between ω_r^2 and ω_0^2 could be sensibly neglected. To my knowledge the work of Schaefer represents the earliest experimental demonstration of the existence of localized vibration modes.

A typical transmission curve obtained by Schaefer, for U-centers in KCl , is shown in Fig. VIII, 1. The pronounced dip in the transmission at a wavelength of about 20μ is due to the strong absorption by the localized vibration mode.

In the frequency interval $(0, \omega_L)$ it can happen that if the function $\alpha(\omega^2)$ is a sufficiently rapidly changing function of ω , then for either $0 < \epsilon < 1$ or $\epsilon < 0$ the equation

$$\omega_j^2(\omega) - \omega^2 + \epsilon p \mu \omega_j^2(\omega) \alpha(\omega^2) = 0 \quad (8. 3. 28)$$

can have a solution, which we denote by ω_r^2 , and in the neighborhood of this frequency the absorption coefficient takes the form

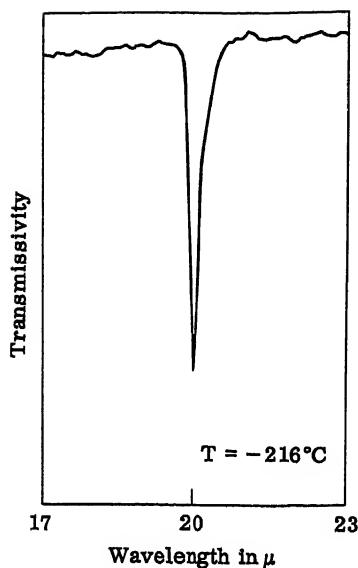


Fig. VIII.1 A transmission spectrum for KCl containing $\sim 10^{17}$ U-centers per cubic centimeter. The strong absorption at a wavelength of 20μ is the absorption resulting from the localized vibration mode [see Eq. (8.3.27)].

$$\alpha(\omega) = \frac{4\pi\omega_r^3}{3c} \frac{M^2\left(\begin{smallmatrix} o \\ j \end{smallmatrix}\right)}{v_a} \frac{1}{[\alpha'(\omega_r^2)]^2} \\ \times \frac{\epsilon\mu\beta(\omega_r^2)}{(\omega^2 - \omega_r^2)^2 + \frac{\epsilon^2 p^2 \mu^2 \omega_j^4(o) \beta^2(\omega_r^2)}{[\alpha'(\omega_r^2)]^2}} \quad (8.3.29)$$

However, in the absence of either experimental or computational verification for the existence of such in-band resonances in the infrared absorption spectrum of mixed or isotopically disordered ionic crystals, the

preceding remarks must be considered only as a conjecture.

Numerical calculations of the absorption spectrum as a function of ω on the basis of the results of the present section would require the use of a high-speed computer. The general features of these spectra, however, are clear from the present analysis.

IX. THE MÖSSBAUER EFFECT

1. The Absorption Cross Section

The absorption cross section per nucleus for the absorption of a γ -ray of energy E by a nucleus in a crystal is given by⁵⁸

$$\sigma_a(E) = \frac{\sigma_0 \Gamma^2}{4} \sum_{ij} \frac{e^{-\beta E_i}}{Z} \frac{\langle i | e^{-i \frac{\mathbf{p}}{\hbar} \cdot \mathbf{R}(t)} | j \rangle \langle j | e^{i \frac{\mathbf{p}}{\hbar} \cdot \mathbf{R}(t)} | i \rangle}{(E_0 - E + E_j - E_i)^2 + \frac{1}{4} \Gamma^2} \quad (9.1.1)$$

where E_0 is the energy difference between the final and initial nuclear states of the absorbing nucleus, E_i and E_j are the energies of the eigenstates $|i\rangle$ and $|j\rangle$ of the crystal, respectively, Γ is the natural width of the excited state of the nucleus, \mathbf{p} is the momentum of the γ ray, $\mathbf{R}(t)$ is the instantaneous position vector of the absorbing nucleus, Z is the crystal's partition function, and σ_0 is the resonance absorption cross section for the nucleus.

If we make use of the result

$$\frac{2a}{a^2 + b^2} = \int_{-\infty}^{\infty} dt e^{-a|t| - ibt}$$

(9.1.2)

we can rewrite Eq. (9.1.1) as

$$\sigma_a(E) = \sigma_o \frac{\Gamma}{4} \sum_{ij} \frac{e^{-\beta \epsilon_i}}{Z} \int_{-\infty}^{\infty} dt e^{-\frac{\Gamma}{2} |t| - i t (E_o - E + E_j - E_i)} \times \langle i | e^{-2\pi i \kappa \cdot R(t)} | j \rangle \langle j | e^{2\pi i \kappa \cdot R(t)} | i \rangle \quad (9.1.3)$$

For convenience we have put

$$\frac{p}{\hbar} = 2\pi\kappa \quad (9.1.4)$$

so that κ is now a vector whose magnitude is the reciprocal of the γ -ray's wavelength and whose direction is that of the γ ray.

For a Bravais lattice, which is all that we consider in this section, the position vector $R(t)$ can be written

$$R(t) = x(t) + u(t) \quad (9.1.5)$$

where $x(t)$ is the position vector of the mean position of the absorbing nucleus, and $u(t)$ is its displacement from the mean position. We can manipulate Eq. (9.1.3) into the form

$$\sigma_a(E) = \sigma_o \frac{\Gamma}{4\hbar} \sum_i \frac{e^{-\beta \epsilon_i}}{Z} \int_{-\infty}^{\infty} dt e^{-\frac{\Gamma}{2\hbar} |t| - i \frac{t}{\hbar} (E_o - E)} \times \langle i | e^{-2\pi i \kappa \cdot u(t;t)} e^{2\pi i \kappa \cdot u(t;o)} | i \rangle \\ = \sigma_o \frac{\gamma}{2} \int_{-\infty}^{\infty} dt e^{i\omega t - \gamma |t|} \langle e^{-2\pi i \kappa \cdot u(t;t)} e^{2\pi i \kappa \cdot u(t;o)} \rangle \quad (9.1.6)$$

where, for convenience, we have introduced the notation

$$\hbar \omega = E - E_0 \quad \gamma = \frac{\Gamma}{2 \hbar} \quad (9.1.7)$$

and $u(\ell; t)$ denotes the Heisenberg operator

$$u(\ell; t) = e^{i \frac{t}{\hbar} H} u(\ell; 0) e^{-i \frac{t}{\hbar} H} \quad (9.1.8)$$

We can simplify Eq. (9.1.7) somewhat, since we are dealing with a harmonic crystal. According to the Baker-Hausdorff formula,⁵⁹ if A and B are two noncommuting operators that commute with their commutator, then

$$e^A e^B = e^{A+B + \frac{1}{2} [A, B]} \quad (9.1.9)$$

For harmonic crystals the commutator $[\kappa \cdot u(\ell, t), \kappa \cdot u(\ell; 0)]$ is a c -number, so that we can write

$$\begin{aligned} e^{-2\pi i \kappa \cdot u(\ell; t)} e^{2\pi i \kappa \cdot u(\ell; 0)} &= e^{-2\pi i \kappa \cdot (u(\ell; t) - u(\ell; 0))} \\ &\times e^{2\pi^2 [\kappa \cdot u(\ell; t), \kappa \cdot u(\ell; 0)]} \end{aligned} \quad (9.1.10)$$

Again, because our crystal is harmonic, we have⁶⁰

$$\langle e^{-2\pi i \kappa \cdot (u(\ell; t) - u(\ell; 0))} \rangle = e^{-2\pi^2 \langle (\kappa \cdot (u(\ell; t) - u(\ell; 0)))^2 \rangle}$$

We have finally the result

$$\begin{aligned}
 & \langle e^{-2\pi i \kappa \cdot u(t;t)} e^{2\pi i \kappa \cdot u(t;o)} \rangle \\
 &= e^{-2\pi^2 \langle (\kappa \cdot u(t;t))^2 \rangle - 2\pi^2 \langle (\kappa \cdot u(t;o))^2 \rangle} \\
 & \quad + 4\pi^2 \langle \kappa \cdot u(t;t) \kappa \cdot u(t;o) \rangle
 \end{aligned}$$

In a time independent system, $\langle (\kappa \cdot u(t;t))^2 \rangle$ is independent of time, so that we can combine the first two terms in the exponential. The absorption cross section can thus be expressed

$$\begin{aligned}
 \sigma_a(E) &= \sigma_o \frac{\gamma}{2} e^{-4\pi^2 \langle (\kappa \cdot u(t))^2 \rangle} \int_{-\infty}^{\infty} dt e^{i\omega t - \gamma|t|} \\
 & \quad \times e^{4\pi^2 \langle \kappa \cdot u(t;t) \kappa \cdot u(t;o) \rangle}
 \end{aligned}
 \tag{9.1.11}$$

If we expand the second exponential factor in the integrand of Eq. (9.1.11) in powers of its argument, the term containing

$1/n! \langle \kappa \cdot u(t;t) \kappa \cdot u(t;o) \rangle^n$ gives the absorption cross section for all processes in which the γ ray is absorbed with the emission of n phonons. We consider only the first two terms in this expansion

$$\sigma_a(E) = \sigma_a^{(0)}(E) + \sigma_a^{(1)}(E) + \dots
 \tag{9.1.12}$$

$$\sigma_a^{(0)}(E) = \sigma_o \frac{\gamma}{2} e^{-4\pi^2 \langle (\kappa \cdot u(t))^2 \rangle} \int_{-\infty}^{\infty} dt e^{i\omega t - \gamma|t|}$$

$$= \sigma_0 \gamma^2 e^{-4\pi^2 \langle (\kappa \cdot u(t))^2 \rangle} \frac{1}{\gamma^2 + \omega^2} \quad (9.1.13)$$

$$\begin{aligned} \sigma_a^{(1)}(E) &= 2\pi^2 \sigma_0 \gamma e^{-4\pi^2 \langle (\kappa \cdot u(t))^2 \rangle} \\ &\times \int_{-\infty}^{\infty} dt e^{i\omega t - \gamma|t|} \langle \kappa \cdot u(t; t) \kappa \cdot u(t; 0) \rangle \end{aligned} \quad (9.1.14)$$

The term $\sigma_a^{(0)}(E)$ gives the cross section for the recoilless absorption of γ rays (the Mössbauer effect)³¹ and is of the form of a very sharp Lorentzian peak of width Γ centered at $E = E_0$. This peak is modulated by the function

$$e^{-4\pi^2 \langle (\kappa \cdot u(t))^2 \rangle} = e^{-2M} \quad (9.1.15)$$

which is sometimes called the resonant fraction, or more commonly, the Debye-Waller factor, after the two physicists who first achieved its evaluation.⁶¹

In deriving Eqs. (9.1.13) and (9.1.14) the only approximation we have made is that the crystal is harmonic. Thus these results hold in the practically important case when the absorbing nucleus is an impurity in the crystal, provided that the approximation of harmonic interatomic forces is still valid for the perturbed crystal. This is the situation which will concern us in what follows.

It is convenient to transform the expression for $\sigma_a^{(1)}(E)$ given by Eq. (9.1.14) into a form more suited for its computation. Let us denote by $B_{xy}(\omega)$ the integral

$$B_{xy}(\ell; \omega) = \int_{-\infty}^{\infty} dt e^{i\omega t - \gamma|t|} \langle u_x(\ell; t) u_y(\ell; 0) \rangle \quad (9.1.16)$$

In terms of this function we have

$$\sigma_a^{(1)}(E) = 2\pi^2 \sigma_0 \gamma e^{-2M} \sum_{xy} \kappa_x B_{xy}(\ell; \omega) \kappa_y \quad (9.1.17)$$

If we express $B_{xy}(\ell; \omega)$ in terms of the eigenstates of the crystal Hamiltonian, we can write it as

$$B_{xy}(\ell; \omega) = \int_{-\infty}^{\infty} d\nu \rho(\nu) \frac{2\gamma}{\gamma^2 + (\omega + \nu)^2} \quad (9.1.18)$$

where the spectral function $\rho(\nu)$ is given by

$$\begin{aligned} \rho(\nu) = & \frac{1}{Z} \sum_{mn} e^{-\beta E_m} \langle m | u_x(\ell; 0) | n \rangle \langle n | u_y(\ell; 0) | m \rangle \\ & \times \delta\left(\nu - \frac{1}{\hbar} (E_m - E_n)\right) \end{aligned} \quad (9.1.19)$$

In their discussion of the Mössbauer effect, Singwi and Sjölander⁶² have made an approximation equivalent to letting $\gamma \rightarrow 0 + i\epsilon$ in Eq. (9.1.18). If we also make this approximation we find that

$$\begin{aligned}
 B_{xy}(\ell; \omega) &= 2\pi \rho(-\omega) \\
 &= 2\pi e^{\beta \hbar \omega} \rho(\omega)
 \end{aligned}
 \tag{9.1.20}$$

where the step connecting the last two expressions is readily taken with the aid of Eq. (9.1.19).

Whether we choose to use Eq. (9.1.18) or (9.1.20), our problem is now to obtain the function $\rho(\nu)$. We do this in such a way that we can utilize directly results we have obtained in Sec. VII.

We begin by introducing the function $f(u)$ which is defined by

$$f(u) = \langle T e^{uH} u_x(\ell; 0) e^{-uH} u_y(\ell; 0) \rangle \tag{9.1.21}$$

This function satisfies the periodicity condition

$$f(u + \beta) = f(u) \quad -\beta < u < 0 \tag{9.1.22}$$

and we, accordingly, expand it in a Fourier series

$$f(u) = \sum_{\ell=-\infty}^{\infty} g_{\ell} e^{\frac{2\pi i \ell u}{\beta}} \tag{9.1.23}$$

The Fourier coefficient g_{ℓ} is given by

$$\begin{aligned}
 g_{\ell} &= \frac{1}{\beta} \int_0^{\beta} f(u) e^{\frac{-2\pi i \ell u}{\beta}} du \\
 &= \frac{1}{\beta \hbar Z} \sum_{mn} e^{-\beta E_m} \langle m | u_x(\ell; 0) | n \rangle \langle n | u_y(\ell; 0) | m \rangle \\
 &\quad \times \frac{e^{\beta(E_m - E_n)} - 1}{\frac{1}{\hbar}(E_m - E_n) - \frac{2\pi i \ell}{\beta \hbar}}
 \end{aligned}
 \tag{9.1.24}$$

We now introduce a function of a continuous variable ν , $g(\nu)$, which equals g_ℓ when ν equals ν_ℓ , where

$$\nu_\ell = \frac{2\pi i \ell}{\beta \hbar} \quad (9.1.25)$$

and $g(\nu)$ is given by

$$g(\nu) = \frac{1}{\beta \hbar Z} \sum_{mn} e^{-\beta E_m} \langle m | u_x(t;0) | n \rangle \langle n | u_y(t;0) | m \rangle \times \frac{e^{\beta(E_m - E_n)} - 1}{\frac{1}{\hbar}(E_m - E_n) - \nu} \quad (9.1.26)$$

It follows that

$$\rho(\nu) = \frac{\beta \hbar}{e^{\beta \hbar \nu} - 1} \lim_{\delta \rightarrow 0+} \frac{g(\nu + i\delta) - g(\nu - i\delta)}{2\pi i} \quad (9.1.27)$$

In the present case $f(u)$ is given explicitly by

$$f(u) = \frac{\hbar}{2NM} \sum_{\substack{k, k' \\ j, j'}} \frac{e_x^{(k)}(j) e_y^{(k')}(j')}{\sqrt{\omega_j(k) \omega_{j'}(k')}} e^{2\pi i (k - k') \cdot x(t)} \langle T (a_{-kj}^+(u) + a_{kj}(u)) (a_{-k'j'}(0) + a_{k'j'}^+(0)) \rangle \quad (9.1.28)$$

and the evaluation of

$$\langle T (a_{-kj}^+(u) + a_{kj}(u)) (a_{-k'j'}(0) + a_{k'j'}^+(0)) \rangle$$

has been described in Sec. VII.

2. The Impurity Debye-Waller Factor

We first discuss the evaluation of the function $2M$ that appears in the exponent of the Debye-Waller factor. We restrict ourselves to the case of a cubic crystal. This function is given by

$$\begin{aligned}
 2M &= 4\pi^2 \langle (\kappa \cdot u(t))^2 \rangle \\
 &= \frac{4\pi^2 \hbar}{2NM} \sum_{\mathbf{k}, \mathbf{k}'} \sum_{\mathbf{j}, \mathbf{j}'} \frac{(\kappa \cdot \mathbf{e}(\mathbf{j})) (\kappa \cdot \mathbf{e}(\mathbf{j}'))}{\sqrt{\omega_{\mathbf{j}}(\mathbf{k}) \omega_{\mathbf{j}'}(\mathbf{k}')}} e^{2\pi i (\mathbf{k} - \mathbf{k}') \cdot \mathbf{x}(t)} \\
 &\quad \times \langle (a_{-\mathbf{k}\mathbf{j}}^+ (0) + a_{\mathbf{k}\mathbf{j}} (0)) (a_{-\mathbf{k}'\mathbf{j}'} (0) + a_{\mathbf{k}'\mathbf{j}'}^+ (0)) \rangle
 \end{aligned} \tag{9.2.1}$$

Let us assume that the absorbing nucleus is an impurity in the host crystal and can be regarded as a substitutional mass defect in which the normal atom of mass M at the site $\mathbf{x}(t)$ has been replaced by an atom of mass $(1 - \epsilon)M$. With no loss of generality we can assume that the absorbing nucleus is located at the origin of our coordinate system, i.e., at $\mathbf{x}(t) = 0$. This is convenient because the calculations of Sec. VII-2 are based on this assumption.

We make use of the results given at the end of Sec. VII-1 together with Eq. (7.2.14) to rewrite Eq. (9.2.1) as

$$\begin{aligned}
 2M &= \frac{4\pi^2 \hbar}{2NM} \sum_{\mathbf{j}, \mathbf{j}'} \sum_{\mathbf{k}, \mathbf{k}'} \sum_{\ell=-\infty}^{\infty} \frac{(\kappa \cdot \mathbf{e}(\mathbf{j})) (\kappa \cdot \mathbf{e}(\mathbf{j}'))}{\sqrt{\omega_{\mathbf{j}}(\mathbf{k}) \omega_{\mathbf{j}'}(\mathbf{k}')}} \\
 &\quad \times \left\{ \frac{2\omega_{\mathbf{j}}(\mathbf{k})}{\beta \hbar} \frac{\Delta(\mathbf{k} - \mathbf{k}') \delta_{\mathbf{j}\mathbf{j}'}}{\omega_{\mathbf{j}}^2(\mathbf{k}) + \omega_{\ell}^2} + \frac{2\omega_{\ell}}{\beta \hbar} \frac{1}{\omega_{\mathbf{j}}^2(\mathbf{k}) + \omega_{\ell}^2} \right. \\
 &\quad \times \left. \frac{\lambda \beta \hbar}{2N} \frac{e(\mathbf{j}) \cdot \mathbf{e}(\mathbf{j}'))}{\sqrt{\omega_{\mathbf{j}}(\mathbf{k}) \omega_{\mathbf{j}'}(\mathbf{k}'))}} \frac{1}{1 + \lambda g_{\ell}(\omega_{\ell}^2)} \frac{2\omega_{\ell}}{\beta \hbar} \frac{1}{\omega_{\mathbf{j}'}^2(\mathbf{k}') + \omega_{\ell}^2} \right\}
 \end{aligned} \tag{9.2.2}$$

The first term is just the result for $2M$ if the absorbing nucleus is one of the normal atoms of the crystal. If we denote it by $2M_0$ we find

$$\begin{aligned}
 2M_0 &= \frac{4\pi^2\hbar}{2NM} \sum_{\mathbf{k}} \frac{(\kappa \cdot \mathbf{e}(\mathbf{j}^{\mathbf{k}}))^2}{\omega_{\mathbf{j}}(\mathbf{k})} \sum_{\ell=-\infty}^{\infty} \frac{2\omega_{\mathbf{j}}(\mathbf{k})}{\beta\hbar} \frac{1}{\omega_{\mathbf{j}}(\mathbf{k}) + \omega_{\ell}} \\
 &= \frac{4\pi^2\hbar}{2NM} \sum_{\mathbf{k}} \frac{(\kappa \cdot \mathbf{e}(\mathbf{j}^{\mathbf{k}}))^2}{\omega_{\mathbf{j}}(\mathbf{k})} [2n_{\mathbf{j}}(\mathbf{k}) + 1]
 \end{aligned}
 \tag{9.2.3}$$

The sum over ℓ in this expression is a standard sum, but can readily be evaluated by contour integration. Since we are dealing with a cubic crystal, this term can be rewritten

$$2M_0 = \frac{(2\pi\kappa)^2\hbar}{6NM} \sum_{\mathbf{k}} \frac{2n_{\mathbf{j}}(\mathbf{k}) + 1}{\omega_{\mathbf{j}}(\mathbf{k})}
 \tag{9.2.4}$$

The second term of Eq. (9.2.2), which we denote by $2M_1$, can also be simplified, if we invoke cubic symmetry. We write this term as

$$\begin{aligned}
 2M_1 &= \frac{4\pi^2\hbar}{2NM} \frac{\lambda\beta\hbar}{2N} \frac{4}{\beta^2\hbar^2} \sum_{\ell=-\infty}^{\infty} \sum_{\alpha\beta\gamma} \\
 &\quad \times \kappa_{\alpha}\kappa_{\beta}e_{\alpha}(\mathbf{j}^{\mathbf{k}})e_{\beta}(\mathbf{j}^{\mathbf{k}'})e_{\gamma}(\mathbf{j}^{\mathbf{k}})e_{\gamma}(\mathbf{j}^{\mathbf{k}'})
 \end{aligned}$$

$$\times \left\{ \frac{\omega_l^2}{\omega_l^2 + \omega_j^2(k)} \frac{1}{1 + \lambda g_l} \frac{\omega_l^2}{\omega_l^2 + \omega_{j'}^2(k')} \right\} \quad (9.2.5)$$

The expression in curly brackets is a function of (k, j) and (k', j') only through the frequencies $\omega_j(k)$ and $\omega_{j'}(k')$, respectively. It therefore is invariant against all the operations which take a cubic crystal into itself. Consequently, $2M_1$ will vanish unless $\alpha = \gamma$ and $\beta = \gamma$.

We can argue further that insofar as the sums over (k, j) and (k', j') are concerned, γ can take on any of the values x, y, z , without altering their values. We therefore obtain the result

$$2M_1 = \lambda \frac{(2\pi\kappa)^2}{\beta M} \sum_{l=-\infty}^{\infty} \frac{\omega_l^2}{1 + \lambda g_l} \left\{ \frac{1}{3N} \sum_k \frac{1}{\omega_l^2 + \omega_j^2(k)} \right\}^2 \quad (9.2.6)$$

To obtain a result that is correct to lowest order in λ , we can set the factor $[1 + \lambda g_l]^{-1}$ equal to unity in Eq. (9.2.6), and we obtain

$$2M_1 \approx \frac{\lambda (2\pi\kappa)^2}{\beta M} \frac{1}{(3N)^2} \sum_{j, j'} \frac{1}{\omega_j^2(k) - \omega_{j'}^2(k')} \\ \times \sum_{l=-\infty}^{\infty} \left\{ \frac{\omega_j^2(k)}{\omega_l^2 + \omega_j^2(k)} - \frac{\omega_{j'}^2(k')}{\omega_l^2 + \omega_{j'}^2(k')} \right\}$$

$$= \lambda \frac{(2\pi\kappa)^2}{2M} \frac{\hbar}{(3N)^2} \sum_{\substack{\mathbf{k} \mathbf{k}' \\ \mathbf{j} \mathbf{j}'}} \frac{1}{\omega_{\mathbf{j}}^2(\mathbf{k}) - \omega_{\mathbf{j}'}^2(\mathbf{k}')} \left\{ \omega_{\mathbf{j}}(\mathbf{k}) [2n_{\mathbf{j}}(\mathbf{k}) + 1] \right. \\ \left. - \omega_{\mathbf{j}'}(\mathbf{k}') [2n_{\mathbf{j}'}(\mathbf{k}') + 1] \right\}$$

(9.2.7)

This result takes simple forms at the absolute zero of temperature and in the high-temperature limit:

$$2M_1 \underset{T=0^\circ\text{K}}{\approx} \lambda \frac{(2\pi\kappa)^2}{2M} \frac{\hbar}{(3N)^2} \sum_{\substack{\mathbf{k} \mathbf{k}' \\ \mathbf{j} \mathbf{j}'}} \frac{1}{\omega_{\mathbf{j}}(\mathbf{k}) + \omega_{\mathbf{j}'}(\mathbf{k}')}$$

(9.2.8a)

$$\underset{T \rightarrow \infty}{\approx} \lambda \frac{(2\pi\kappa)^2 \hbar^2}{12 M k T}$$

(9.2.8b)

The latter result has recently been obtained by Maradudin and Flinn⁶³ by a different method.

In order to obtain a nonperturbation result we have to go back to Eq. (9.2.6) and adopt some model for the host crystal. If we assume a Debye spectrum, we find that

$$\frac{1}{3N} \sum_{\mathbf{j}} \frac{1}{\omega_{\mathbf{l}}^2 + \omega_{\mathbf{j}}^2(\mathbf{k})} = \frac{3}{\omega_{\text{D}}^2} \left[1 - \xi_{\mathbf{l}} \tan^{-1} \frac{1}{\xi_{\mathbf{l}}} \right]$$

(9.2.9a)

$$g_{\mathbf{l}} = 1 - \frac{\omega_{\mathbf{l}}^2}{3N} \sum_{\mathbf{j}} \frac{1}{\omega_{\mathbf{l}}^2 + \omega_{\mathbf{j}}^2(\mathbf{k})} = 1 - 3 \xi_{\mathbf{l}}^2 \left[1 - \xi_{\mathbf{l}} \tan^{-1} \frac{1}{\xi_{\mathbf{l}}} \right]$$

(9.2.9b)

where

$$\xi_l = \frac{\omega_l}{\omega_D} = \frac{2\pi l}{\beta \hbar \omega_D} \quad (9.2.10)$$

The expression for $2M_1$ now becomes

$$2M_1 = 9\epsilon \frac{(2\pi\kappa)^2}{\beta M \omega_D^2} \sum_{l=-\infty}^{\infty} \frac{\xi_l^2}{1 - 3\epsilon \xi_l^2 (1 - \xi_l \tan^{-1} \frac{1}{\xi_l})} \times [1 - \xi_l \tan^{-1} \frac{1}{\xi_l}] \quad (9.2.11)$$

At the absolute zero of temperature we can replace summation over l by integration over ξ according to

$$\sum_{l=-\infty}^{\infty} \rightarrow \frac{\beta \hbar \omega_D}{2\pi} \int_{-\infty}^{\infty} d\xi \quad (9.2.12)$$

and we obtain

$$2M_1 =_{0^\circ K} 9\epsilon \frac{(2\pi\kappa)^2}{M \omega_D^2} \frac{\hbar}{2\pi} \int_{-\infty}^{\infty} \frac{\xi^2 (1 - \xi \tan^{-1} \frac{1}{\xi})^2 d\xi}{1 - 3\epsilon \xi^2 (-\xi \tan^{-1} \frac{1}{\xi})} \quad (9.2.13)$$

The integral can be evaluated numerically for different values of ϵ .

At high temperatures we seem to be limited to expanding the summand

in Eq. (9.2.11) in inverse powers of ξ_l :

$$\begin{aligned}
 2 M_1 &= \frac{18 \epsilon}{1 - \epsilon} \frac{(2 \pi \kappa)^2}{\beta M \omega_D^2} \sum_{l=1}^{\infty} \left\{ \frac{1}{9 \xi_l^2} - \frac{1}{15} \frac{2 - \epsilon}{1 - \epsilon} \frac{1}{\xi_l^4} + \dots \right\} \\
 &= \frac{18 \epsilon}{1 - \epsilon} (2 \pi \kappa)^2 \frac{k T}{M \omega_D^2} \\
 &\times \left\{ \frac{1}{9} \frac{\pi^2}{6} \left(\frac{\hbar \omega_D}{2 \pi k T} \right)^2 - \frac{1}{15} \frac{2 - \epsilon}{1 - \epsilon} \frac{\pi^4}{90} \left(\frac{\hbar \omega_D}{2 \pi k T} \right)^4 + \dots \right\}
 \end{aligned}
 \tag{9.2.14}$$

3. The One-Phonon Absorption Cross Section

To calculate the one-phonon absorption cross section we adopt the approximation expressed by Eq. (9.1.19). By combining Eqs. (9.1.16), (9.1.19), and (9.1.26), we have

$$\begin{aligned}
 \sigma_a^{(1)}(\omega) &= 2 \pi^2 \sigma_0 \gamma e^{-2 M} \sum_{xy} \kappa_x \kappa_y 2 \pi \beta \hbar \frac{1}{1 - e^{-\beta \hbar \omega}} \\
 &\times \lim_{\delta \rightarrow 0+} \frac{g(\omega + i \delta) - g(\omega - i \delta)}{2 \pi i}
 \end{aligned}
 \tag{9.3.1}$$

The function $g(\omega)$ is obtained from Eq. (9.1.28) and (7.2.14) and is

$$\begin{aligned}
g(\nu) = & \frac{\hbar}{2NM} \sum_{kj} \frac{e_x^{(k)} e_y^{(k)}}{\omega_j(k)} \frac{2\omega_j(k)}{\beta\hbar} \frac{1}{\omega_j^2(k) - \nu^2} \\
& - \frac{\hbar}{2NM} \frac{4\nu^2}{\beta^2\hbar^2} \frac{\epsilon\beta\hbar}{2N} \sum_{\substack{k,k' \\ j,j'}} e_x^{(k)} e_y^{(k')} \\
& \times \frac{1}{\omega_j^2(k) - \nu^2} \frac{1}{\omega_{j'}^2(k') - \nu^2} \frac{e^{(k)} : e^{(k')}}{1 - \epsilon\nu^2 \frac{1}{3N} \sum_{k_1 j_1} \frac{1}{\nu^2 - \omega_{j_1}^2(k_1)}}
\end{aligned}
\tag{9.3.2}$$

We consider the contributions to $\sigma_a^{(1)}(E)$ from each of the two terms in Eq. (9.3.2) separately.

The contribution to $\sigma_a^{(1)}(E)$ from the first term of Eq. (9.3.2) is

$$\begin{aligned}
\sigma_a^{(10)}(E) = & 4\pi^3 \sigma_o \gamma \frac{e^{-2M}}{1 - e^{-\beta\hbar\omega}} \frac{\hbar}{M} \\
& \times \frac{1}{N} \sum_{kj} (\kappa \cdot e^{(k)})^2 \delta(\omega^2 - \omega_j^2(k))
\end{aligned}
\tag{9.3.3}$$

If we use the cubic symmetry of our crystal, this result can be written

$$\sigma_a^{(10)}(E) = (2\pi\kappa)^2 \pi \sigma_0 \gamma \frac{e^{-2M}}{1 - e^{-\beta \hbar \omega}} \frac{\hbar}{M} G_0(\omega^2) \quad (9.3.4)$$

where $G_0(\omega^2)$ is the distribution function for the squares of the normal mode frequencies of the unperturbed crystal.

The result expressed by Eq. (9.3.4) is the basis for the suggestion by Visscher⁶⁴ that the absorption of γ rays by nuclei bound in crystals could be used to obtain the crystal's frequency spectrum.

With the aid of cubic symmetry we can rewrite the second term of Eq. (9.3.2) as

$$\begin{aligned} & - \frac{\epsilon \nu^2}{\beta M} \frac{\delta_{xy}}{(3N)^2} \sum_{\substack{\mathbf{k} \mathbf{k}' \\ j j'}} \frac{1}{\omega_j^2(\mathbf{k}) - \nu^2} \frac{1}{\omega_{j'}^2(\mathbf{k}') - \nu^2} \\ & \times \frac{1}{1 - \epsilon \nu^2 \frac{1}{3N} \sum_{\substack{\mathbf{k}_1 \\ j_1}} \frac{1}{\nu^2 - \omega_{j_1}^2(\mathbf{k}_1)}} \end{aligned} \quad (9.3.5)$$

We first consider the case $0 < \epsilon < 1$ and $0 < \omega < \omega_L$. By substituting Eq. (9.3.5) into Eq. (9.3.1), we obtain

$$\sigma_a^{(11)}(E) = - (2\pi\kappa)^2 \sigma_0 \gamma \epsilon \frac{\hbar}{M} \frac{e^{-2M}}{1 - e^{-\beta \hbar \omega}}$$

$$\times \text{Im} \left\{ -\tilde{G}_0(\omega^2) + i\pi G_0(\omega^2) \right\}^2$$

$$\times \frac{\omega^2 [A(\omega^2) - i B(\omega^2)]}{A^2(\omega^2) + B^2(\omega^2)} \quad (9.3.6)$$

where

$$A(\omega^2) = 1 - \frac{\epsilon \omega^2}{3N} \sum_{\mathbf{kj}} \frac{1}{(\omega^2 - \omega_j^2(\mathbf{k}))^2} = 1 - \epsilon \omega^2 \tilde{G}_O(\omega^2) \quad (9.3.7a)$$

$$B(\omega^2) = \epsilon \pi \omega^2 G_O(\omega^2) \quad (9.3.7b)$$

We thus find

$$\begin{aligned} \sigma_a^{(11)}(E) &= (2\pi\kappa)^2 \sigma_O \gamma \frac{\epsilon \hbar \omega^2}{M} \frac{e^{-2M}}{1 - e^{-\beta \hbar \omega}} \frac{1}{\epsilon^2 \omega^4} \frac{B}{A^2 + B^2} \\ &\quad - (2\pi\kappa)^2 \pi \sigma_O \gamma \frac{e^{-2M}}{1 - e^{-\beta \hbar \omega}} \frac{\hbar}{M} G_O(\omega^2) \end{aligned} \quad (9.3.8)$$

By combining Eqs. (9.3.4) and (9.3.8), we obtain for the one-phonon absorption cross section

$$\sigma_a^{(1)}(E) = (2\pi\kappa)^2 \pi \sigma_O \gamma \frac{e^{-2M}}{1 - e^{-\beta \hbar \omega}} \frac{\hbar}{M} G_O(\omega^2)$$

$$\times \frac{1}{\left[1 - \epsilon \omega^2 \tilde{G}_0(\omega^2)\right]^2 + \epsilon^2 \pi^2 \omega^4 G_0^2(\omega^2)} \quad (9.3.9)$$

a remarkably simple expression.

We turn now to the case that $0 < \epsilon < 1$ and $\omega > \omega_L$. Unless ϵ is sufficiently large, there is no contribution to the one-phonon absorption cross section. However, if ϵ is large enough that the equation

$$1 = \epsilon \omega^2 \frac{1}{3N} \sum_{\mathbf{k}} \frac{1}{\omega^2 - \omega_j^2(\mathbf{k})} \quad (9.3.10)$$

has a solution, which we denote by ω_0 , then we obtain as the only contribution to the one-phonon absorption cross section

$$\begin{aligned} \sigma_a^{(11)}(E) = & (2\pi\hbar^2)^2 \pi \sigma_0 \gamma \frac{\hbar}{M} \frac{e^{-2M}}{1 - e^{-\beta\hbar\omega}} [\tilde{G}_0(\omega_0^2)]^2 \\ & \times \frac{\delta(\omega^2 - \omega_0^2) \operatorname{sgn} \omega}{\frac{d}{d\omega^2} \left\{ -\omega^2 \tilde{G}_0(\omega^2) \right\}_{\omega^2 = \omega_0^2}} \end{aligned} \quad (9.3.11)$$

We see that, if the absorbing nucleus is sufficiently light so that it gives rise to a localized vibration mode whose frequency is ω_0 , then the absorption cross section will show a very sharp peak at a γ -ray energy given by $\hbar\omega_0 = E - E_0$. The suggestion that the Mössbauer effect be used to show up the existence of a localized vibration mode has been made by several authors.⁶⁵

We finally come to the case $\epsilon < 0$. In this case there is no

contribution to the absorption spectrum for $\omega > \omega_L$. In the interval $(0, \omega_L)$ the impurity contribution to the one-phonon absorption cross section is still given by Eq. (9.3.9) except that ϵ is now replaced by $-\epsilon$.

Returning now to Eq. (9.3.9), we can use the results of Nardelli and Tettamanzi¹⁷ to make some interesting predictions about the behavior of $\sigma_a^{(1)}(\omega)$ as a function of ω . These authors have plotted $\omega^2 \tilde{G}_0(\omega^2)$ as a function of ω^2 for a nearest neighbor, central force model of a face-centered cubic crystal. This plot is reproduced (in our notation) in Fig. IX.1. (Recall that because of difference in normalization, the Nardelli and Tettamanzi function $1/3 \omega^2 \tilde{G}_0(\omega^2)$ equals our $\tilde{G}_0(\omega^2)$.)

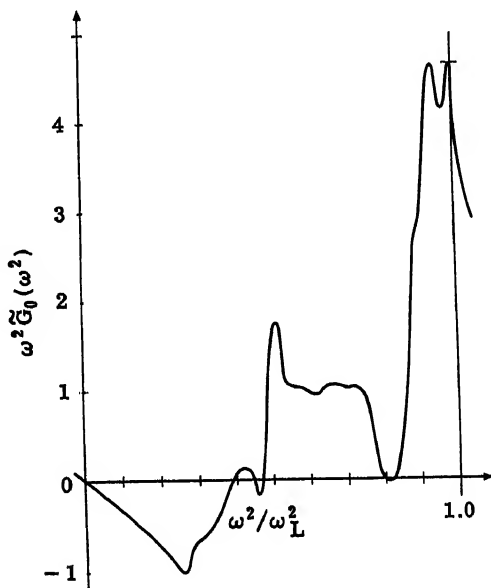


Fig. IX.1. A plot of the function $\omega^2 \tilde{G}_0(\omega^2)$ for a face-centered cubic crystal with nearest neighbor, central force interactions.

From this plot we see that, depending on the value of ϵ , for $1/\epsilon$ in the interval $-1 < 1/\epsilon < 4.6$ the equation

$$1 - \epsilon \omega^2 \tilde{G}_0(\omega^2) = 0 \quad (9.3.12)$$

can have from one to seven solutions. In the neighborhood of one of these solutions, which we denote by ω_r^2 , $\sigma_a^{(1)}(E)$ has the form

$$\begin{aligned} \sigma_a^{(1)}(E) \approx & (2\pi\kappa)^2 \pi \sigma_0 \gamma \frac{e^{-2M}}{1 - e^{-\beta \hbar \omega}} \frac{\hbar}{M} G_0(\omega_r^2) \\ & \times \frac{1}{4\omega_r^2 A^2(\omega_r^2)} \frac{1}{(\omega - \omega_r)^2 + \Gamma^2(\omega_r^2)} \end{aligned} \quad (9.3.13)$$

where

$$\Gamma(\omega_r^2) = \left| \frac{\epsilon \pi \omega_r G_0(\omega_r^2)}{2 A(\omega_r^2)} \right| \quad (9.3.14)$$

$$A(\omega_r^2) = \frac{d}{d\omega^2} \left\{ 1 - \epsilon \omega^2 \tilde{G}_0(\omega^2) \right\}_{\omega^2 = \omega_r^2} \quad (9.5.15)$$

The ratio of the height of this resonance peak relative to the height of the one-phonon spectrum at $\omega = \omega_r$, which would exist if the absorbing nucleus were one of the atoms of the host crystal and not an impurity, is

$$\text{height} = \frac{1}{\epsilon^2 \pi^2 \omega_r^4 G_0^2(\omega_r^2)} \quad (9.3.16)$$

This quantity should be large if the resonance is to be observed. The width of the peak at half-maximum is

$$\text{width} = 2 \Gamma(\omega_r^2) = \left| \frac{\epsilon \pi \omega_r G_0(\omega_r^2)}{A(\omega_r^2)} \right| \quad (9.3.17)$$

This width should be small if the resonance is to be observed by the Mössbauer effect.

We can make an estimate of the height and width of such a resonance peak if we adopt the Debye spectrum. In this case the functions

$\tilde{G}_0(\omega^2)$ and $G_0(\omega^2)$ are given by

$$\tilde{G}_0(\omega^2) = \frac{3\omega}{2\omega_D^3} \ln \frac{\omega_D + \omega}{\omega_D - \omega} - \frac{3}{\omega_D^2} \quad 0 < \omega < \omega_D \quad (9.3.18a)$$

$$G_0(\omega^2) = \frac{3}{2} \frac{\omega}{\omega_D^3} \quad 0 < \omega < \omega_D \quad (9.3.18b)$$

Let us choose the case that $\epsilon < 0$. The solution of Eq. (9.3.12) which vanishes as $|\epsilon| \rightarrow \infty$ is given by

$$\frac{\omega_r^2}{\omega^2} = \frac{1}{3|\epsilon|} + \frac{1}{9|\epsilon|^2} + \frac{7}{81|\epsilon|^3} + \dots \quad (9.3.19)$$

There is a second solution to Eq. (9.3.12) in the interval $(0, \omega_L)$ but we shall not consider it here. We now assume that $|\epsilon|$ is so large that we can neglect all terms in the expansion (9.3.19) past the first. To this approximation we find that

$$A(\omega_r^2) \approx -\frac{3|\epsilon|}{\omega_D^2} \quad (9.3.20)$$

The results for the height and width become

$$\text{height} \approx \frac{12}{\pi^2} |\epsilon| \quad (9.3.21a)$$

$$\text{width} \approx \frac{\pi}{6} \frac{\omega_D}{|\epsilon|} = \frac{\pi}{2} \left(\frac{\omega_r}{\omega_D} \right) \omega_r \quad (9.3.21b)$$

If we assume that $|\epsilon| = 16$, as would be the case if Sn^{119} were placed in Li^7 , we find

$$\frac{\omega_r}{\omega_D} = 0.14$$

$$\text{height} \approx 20$$

$$\frac{\text{width}}{\omega_r} = 0.226 \quad (9.3.22)$$

The larger $|\epsilon|$ is, the more favorable are the chances of observing this low-frequency resonance with the Mössbauer effect.

The existence of this low-frequency peak in the one-phonon absorption cross section when the absorbing nucleus is a heavy impurity was first pointed out by Brout and Visscher.⁶⁶

Note added in proof:

A more careful analysis of the zero-phonon and one-phonon absorption cross sections than described in this section yields the results that, as they stand, Eqs. (9.1.13) and (9.3.9) through (9.3.11) do not give the complete expressions for these cross sections if the resonant nucleus gives rise to a localized vibration mode. The modifications of the present analysis required to obtain the correct expressions are described, for example, in the second paper in reference 60. The zero-phonon cross section is obtained by multiplying the right side of Eq. (9.1.13) by

$$I_0 \left((2\pi\kappa)^2 \frac{1}{M\epsilon^2} \frac{\hbar z_0}{\omega_0^3 B_0} \right)$$

where $I_n(x)$ is the Bessel function of pure imaginary argument, ω_0 is the local mode frequency, z_0 is the partition function for the local mode os-

cillator, and B_0 is given by

$$B_0 = \frac{1}{3N} \sum_{kj} \frac{\omega_j^2(k)}{(\omega_0^2 - \omega_j^2(k))^2}$$

To obtain the correct one-phonon absorption cross section, the right side of Eq. (9.3.9) must be multiplied by

$$I_0 \left((2\pi\kappa)^2 \frac{1}{M\epsilon^2} \frac{\hbar z_0}{\omega_0^3 B_0} \right)$$

whereas Eq. (9.3.11) should be replaced by

$$\sigma_a^{(1)}(E) = \frac{1}{2} \sigma_0 \gamma e^{-2M} I_1 \left((2\pi\kappa)^2 \frac{1}{M\epsilon^2} \frac{\hbar z_0}{\omega_0^3 B_0} \right) \\ \times \left\{ e^{-\frac{1}{2}\beta\hbar\omega_0} \frac{2\gamma}{\gamma^2 + (\omega + \omega_0)^2} + e^{\frac{1}{2}\beta\hbar\omega_0} \frac{2\gamma}{\gamma^2 + (\omega - \omega_0)^2} \right\}$$

The detailed derivation of these results will be published elsewhere. I should like to acknowledge helpful correspondence with Dr. W. M. Visser on this point.

X. LIFETIME OF A LOCALIZED VIBRATION MODE

1. Introduction

It is by now well known that the substitution of a sufficiently light impurity atom for a normal atom in a crystal gives rise to a normal mode of vibration whose frequency lies above the maximum frequency of the unperturbed crystal. Such a mode, usually called an impurity mode, is further characterized by the fact that the amplitudes of the atoms vibrating in this mode decay exponentially or faster with increasing distance from the impurity site.

If the atoms comprising the crystal interact with strict Hooke's law forces, i. e., if the harmonic approximation is rigorously applicable for describing the atomic vibrations, then the impurity mode is an exact

eigenstate of the crystal Hamiltonian, and is infinitely long-lived. In fact, however, no crystal is perfectly harmonic, and the otherwise independent normal modes of the crystal in the harmonic approximation are coupled by the anharmonic terms in the crystal's potential energy. This coupling, which allows energy to be exchanged among the normal modes of the harmonic crystal, imparts a complex shift to the frequency of each of these modes. The real part of this shift represents an actual change in the value of the frequency, whereas the imaginary part can be interpreted as half the reciprocal of the lifetime of this mode. Both the real and imaginary parts of this complex shift are temperature dependent.

In particular, the frequency of the impurity mode undergoes such a complex shift. The first calculation of the lifetime of an impurity mode was carried out by Klemens⁹ who used perturbation theory. He found that at low temperatures the lifetime is of the order of 100 vibrational periods, and decreases with increasing temperature.

It has been suggested recently that impurity vibrational modes in crystals may be detected experimentally by means of the Mössbauer effect.⁶⁵ Since the feasibility of performing such an experiment depends on the value of the impurity mode lifetime, it was felt that an independent calculation of this quantity by a different method from that employed by Klemens would be worth while.

In this section, we present such a calculation. We derive an expression for the phonon propagator for a crystal possessing cubic anharmonicities into which a light impurity atom has been introduced substitutionally. The poles of this propagator occur at the new normal mode frequencies of the perturbed crystal. The imaginary part of the frequency at which the pole associated with the impurity mode occurs leads to the impurity mode lifetime. The formal expression for the impurity mode lifetime is evaluated for a simple model of a crystal.

The Hamiltonian for an anharmonic Bravais crystal containing a mass defect at the origin is

$$H = \sum_{\alpha} \frac{p_{\alpha}^2(t)}{2M} + \frac{1}{2} \sum_{\substack{\alpha \\ t, t'}} \bar{\Phi}_{\alpha\beta}(t, t') u_{\alpha}(t) u_{\beta}(t') \\ + \sum_{\alpha} p_{\alpha}^2(0) \left(\frac{1}{2M'} - \frac{1}{2M} \right)$$

$$\begin{aligned}
& + \frac{1}{6} \sum_{\substack{\alpha \\ \beta \\ \gamma}} \Phi_{\alpha\beta\gamma}(\alpha\alpha'\alpha'') u_{\alpha}(\alpha) u_{\beta}(\alpha') u_{\gamma}(\alpha'') \\
& + \frac{1}{24} \sum_{\substack{\alpha \\ \beta \\ \gamma \\ \delta}} \Phi_{\alpha\beta\gamma\delta}(\alpha\alpha'\alpha''\alpha''') u_{\alpha}(\alpha) u_{\beta}(\alpha') u_{\gamma}(\alpha'') u_{\delta}(\alpha''') \\
& + \dots
\end{aligned} \tag{10.1.1}$$

$$= H_0 + V + H_A \tag{10.1.2}$$

In Eq. (10.1.1), $\Phi_{\alpha\beta\gamma}(\alpha\alpha'\alpha'')$, $\Phi_{\alpha\beta\gamma\delta}(\alpha\alpha'\alpha''\alpha''')$, \dots are the third-order, fourth-order, etc., atomic force constants of the crystal. They are the partial derivatives of the total potential energy with respect to three, four, etc., displacement amplitudes.

If we express the Hamiltonian in terms of creation and destruction operators, we obtain (again neglecting the zero-point energy)

$$H_0 = \sum_{kj} \hbar\omega_j(k) a_{kj}^+ a_{kj} \tag{10.1.3a}$$

$$\begin{aligned}
V = \sum_{\substack{k_1 k_2 \\ j_1 j_2}} U \left(\begin{matrix} k_1 k_2 \\ j_1 j_2 \end{matrix} \right) & \left(a_{-k_1 j_1}^+ + a_{k_1 j_1} \right) \\
& \left(a_{-k_2 j_2}^+ + a_{k_2 j_2} \right)
\end{aligned} \tag{10.1.3b}$$

$$\begin{aligned}
H_A = & \frac{\hbar^{3/2}}{2^{3/2} 6\sqrt{N}} \sum_{\substack{k_1 k_2 k_3 \\ j_1 j_2 j_3}} \Delta(k_1 + k_2 + k_3) \frac{\mathfrak{E} \begin{pmatrix} k_1 k_2 k_3 \\ j_1 j_2 j_3 \end{pmatrix}}{\sqrt{\omega_{j_1}(k_1) \omega_{j_2}(k_2) \omega_{j_3}(k_3)}} \\
& \times \prod_{p=1}^3 \left(a_{-k_p j_p}^+ + a_{k_p j_p} \right) + \frac{\hbar^2}{2^2 24 N} \\
& \times \sum_{\substack{k_1 k_2 k_3 k_4 \\ j_1 j_2 j_3 j_4}} \Delta(k_1 + k_2 + k_3 + k_4) \frac{\mathfrak{E} \begin{pmatrix} k_1 k_2 k_3 k_4 \\ j_1 j_2 j_3 j_4 \end{pmatrix}}{\sqrt{\omega_{j_1}(k_1) \omega_{j_2}(k_2) \omega_{j_3}(k_3) \omega_{j_4}(k_4)}} \\
& \times \prod_{p=1}^4 \left(a_{-k_p j_p}^+ + a_{k_p j_p} \right) + \dots
\end{aligned} \tag{10.1.3c}$$

where

$$U \begin{pmatrix} k_1 k_2 \\ j_1 j_2 \end{pmatrix} = \lambda \frac{\hbar}{4N} \sqrt{\omega_{j_1}(k_1) \omega_{j_2}(k_2)} e \begin{pmatrix} k_1 \\ j_1 \end{pmatrix} \cdot e \begin{pmatrix} k_2 \\ j_2 \end{pmatrix} \tag{10.1.4a}$$

and

$$\mathfrak{E} \begin{pmatrix} k_1 k_2 k_3 \\ j_1 j_2 j_3 \end{pmatrix} = \sum_{\substack{\alpha \\ j''_1 j''_2 j''_3}} \mathfrak{E}_{\alpha \beta \gamma}(0 \ell' \ell'') \frac{e_{\alpha} \begin{pmatrix} k_1 \\ j''_1 \end{pmatrix} e_{\beta} \begin{pmatrix} k_2 \\ j''_2 \end{pmatrix} e_{\gamma} \begin{pmatrix} k_3 \\ j''_3 \end{pmatrix}}{M^{3/2}}$$

$$\times e^{2\pi i \left(\mathbf{k}_2 \cdot \mathbf{x}(\ell') + \mathbf{k}_3 \cdot \mathbf{x}(\ell'') \right)} \quad (10.1.4b)$$

$$\begin{aligned} \Phi \left(\begin{matrix} k_1 k_2 k_3 k_4 \\ j_1 j_2 j_3 j_4 \end{matrix} \right) &= \sum_{\substack{\alpha \\ \ell' \beta \\ \ell'' \gamma \\ \ell''' \delta}} \Phi_{\alpha \beta \gamma \delta} (0 \ell' \ell'' \ell''') \\ &\times \frac{e_{\alpha} \left(\begin{matrix} k_1 \\ j_1 \end{matrix} \right) e_{\beta} \left(\begin{matrix} k_2 \\ j_2 \end{matrix} \right) e_{\gamma} \left(\begin{matrix} k_3 \\ j_3 \end{matrix} \right) e_{\delta} \left(\begin{matrix} k_4 \\ j_4 \end{matrix} \right)}{M^2} \\ &\times e^{2\pi i \left(\mathbf{k}_2 \cdot \mathbf{x}(\ell') + \mathbf{k}_3 \cdot \mathbf{x}(\ell'') + \mathbf{k}_4 \cdot \mathbf{x}(\ell''') \right)} \end{aligned} \quad (10.1.4c)$$

In writing Eqs. (10.1.4b-c) we have used the fact that, because of the invariance of the crystal against rigid body translation through any translation vector, $\Phi_{\alpha \beta \gamma} \dots (\ell \ell' \ell'' \dots)$ depends only on the differences between ℓ and the remaining position vectors,

$$\Phi_{\alpha \beta \gamma} \dots (\ell \ell' \ell'' \dots) = \Phi_{\alpha \beta \gamma} (0 \ell' - \ell \ell'' - \ell \dots) \quad (10.1.5)$$

The n -th order anharmonic term in the expansion of the potential energy is smaller than the harmonic term by approximately a factor $(u/a_0)^{n-2}$, where u is a root mean square displacement amplitude and a_0 is the lattice parameter.⁶⁷ This fact gives us a convenient way to order the various anharmonic contributions to any physical quantity of interest. We can write H_A formally as

$$H_A = \eta V_3 + \eta^2 V_4 + \dots \quad (10.1.6)$$

where η is an order parameter which can be set equal to unity at the end of the calculation. If we retain terms in H_A up to some particular (even) order in η , we are assured of obtaining a consistent result if we retain terms to the same order in η in our final answer. In what follows we work only to $O(\eta^2)$.

2. The Lifetime of a Localized Vibration Mode

To obtain the lifetime of the localized mode we construct the function

$$F(\omega) = \sum_{kj} \sum_m \frac{e^{-\beta E_m}}{Z} \langle m | \left(a_{-kj}^+ + a_{kj} \right) \times \delta \left(\omega + \frac{E_m}{\hbar} - \frac{H}{\hbar} \right) \left(a_{-kj} + a_{kj}^+ \right) | m \rangle \quad (10.2.1a)$$

$$= \sum_{\substack{k \\ j}} \rho \left(\begin{matrix} k \\ j \end{matrix}; \omega \right) \quad (10.2.1b)$$

Considered as a function of ω , $F(\omega)$ has a δ function peak whenever $a_{kj}^+ |m\rangle$ or $a_{-kj} |m\rangle$ is an eigenstate of the crystal Hamiltonian H . The difference between the energy of the state $a_{kj}^+ |m\rangle$ and the energy of the state $|m\rangle$ can be described as resulting from the addition of one quantum of vibrational energy to the crystal. A similar result holds for the state $a_{-kj} |m\rangle$. In the case of a perfect, harmonic crystal we see that $F(\omega)$ will have singularities at

$\omega = \pm \omega_j(k)$. In the case of an anharmonic crystal these δ -function peaks are replaced by resonances centered at the phonon frequencies and whose widths at half-maximum are the reciprocals of the phonon lifetimes.

Our task, therefore, is to evaluate the function $F(\omega)$ for an anharmonic crystal containing a single mass defect and to determine the width of the resonance associated with the localized mode. We do this by manipulating the expression for $\rho(kj; \omega)$ into a form which we recognize as having studied previously. We have

$$\rho \left(\begin{matrix} k \\ j \end{matrix}; \omega \right) = \frac{1}{Z} \sum_{mn} e^{-\beta E_m} \langle m | a_{-kj}^+ + a_{kj} | n \rangle \times \langle n | a_{-kj} + a_{kj}^+ | m \rangle \delta \left(\omega + \frac{1}{\hbar} (E_m - E_n) \right) \quad (10.2.2)$$

We now introduce the function

$$F \left(\begin{matrix} k \\ j \end{matrix}; u \right) = \langle T e^{uH} (a_{-kj}^+ + a_{kj}) e^{-uH} (a_{-kj} + a_{kj}^+) \rangle \quad (10.2.3)$$

It is a simple matter to show that $F(k, j; u)$ satisfies the condition

$$F \left(\begin{matrix} k \\ j \end{matrix}; u \right) = F \left(\begin{matrix} k \\ j \end{matrix}; u + \beta \right) \quad -\beta < u < 0 \quad (10.2.4)$$

We, accordingly, expand it in a Fourier series

$$F \left(\begin{matrix} k \\ j \end{matrix}; u \right) = \sum_{\ell=-\infty}^{\infty} a_{\ell} \left(\begin{matrix} k \\ j \end{matrix} \right) e^{\frac{2\pi i \ell u}{\beta}} \quad (10.2.5)$$

The Fourier coefficient $a_{\ell}(k, j)$ is given by

$$a_{\ell} \left(\begin{matrix} k \\ j \end{matrix} \right) = \frac{1}{\beta} \int_0^{\beta} F \left(\begin{matrix} k \\ j \end{matrix}; u \right) e^{-\frac{2\pi i \ell u}{\beta}} du$$

$$= \frac{1}{\beta \hbar} \frac{1}{Z} \sum_{mn} e^{-\beta E_m} \langle m | a_{-kj}^+ + a_{kj} | n \rangle$$

$$\times \langle n | a_{-kj} + a_{kj}^+ | m \rangle = \frac{e^{\beta(E_m - E_n)} - 1}{\frac{1}{\hbar} (E_m - E_n) - \frac{2\pi i \ell}{\beta \hbar}}$$

(10.2.6)

We now introduce the function of a continuous variable $a(kj; \nu)$ which equals a_ℓ when

$$\nu = \nu_\ell = \frac{2\pi i \ell}{\beta \hbar} \quad (10.2.7)$$

It is given explicitly by

$$a \left(\begin{matrix} k \\ j \end{matrix}; \nu \right) = \frac{1}{\beta \hbar} \frac{1}{Z} \sum_{mn} e^{-\beta E_m} \langle m | a_{-kj}^+ + a_{kj} | n \rangle$$

$$\times \langle n | a_{-kj} + a_{kj}^+ | m \rangle = \frac{e^{\beta(E_m - E_n)} - 1}{\frac{1}{\hbar} (E_m - E_n) - \nu}$$

(10.2.8)

We therefore see that

$$\rho \left(\begin{matrix} k \\ j \end{matrix}; \omega \right) = \frac{\beta \hbar}{1 - e^{-\beta \hbar \omega}} \lim_{\delta \rightarrow 0^+} \frac{a \left(\begin{matrix} k \\ j \end{matrix}; -\omega - i\delta \right) - a \left(\begin{matrix} k \\ j \end{matrix}; -\omega + i\delta \right)}{2\pi i}$$

(10.2.9)

From its definition, Eq. (10.2.3), we see that $F(kj; u)$ is just a diagonal element of the matrix we have denoted in previous sections by $F(kj; k'j'; u)$. This means that $a_{\ell}(kj)$ is a diagonal element of the matrix we have referred to previously as $A_{\ell}(kj; k'j')$.

In the case of a single mass defect at $x(\ell) = 0$ in a cubic Bravais crystal, we obtain from Eq. (7.2.14) the result that, for a purely harmonic crystal,

$$a_{\ell} \begin{pmatrix} k \\ j \end{pmatrix} = \alpha_{\ell} \begin{pmatrix} k \\ j \end{pmatrix} - \frac{\lambda \beta \hbar \omega_j(k)}{2N} \beta_{\ell}^2 \begin{pmatrix} k \\ j \end{pmatrix} \times \frac{1}{1 + \frac{\lambda \beta \hbar}{2} \frac{1}{3N} \sum_{\substack{k_1 \\ j_1}} \alpha_{\ell} \begin{pmatrix} k_1 \\ j_1 \end{pmatrix} \omega_{j_1}(k_1)} \quad (10.2.10)$$

If we replace $\alpha_{\ell}(kj)$ and $\beta_{\ell}(kj)$ by their explicit expressions, Eq. (7.1.27), and replace $i\omega_{\ell}$ by ν , we find that

$$a \begin{pmatrix} k \\ j \end{pmatrix} \nu = \frac{2\omega_j(k)}{\beta \hbar} \frac{1}{\omega_j^2(k) - \nu^2} - \frac{\epsilon \beta \hbar \omega_j(k)}{2N} \times \left\{ \frac{2\nu}{\beta \hbar} \frac{1}{\omega_j^2(k) - \nu^2} \right\}^2 \left[1 - \frac{\epsilon \nu^2}{3N} \sum_{\substack{k_1 \\ j_1}} \frac{1}{\nu^2 - \omega_{j_1}^2(k_1)} \right]^{-1} \quad (10.2.11)$$

Since we are interested in the local mode frequency, we assume that $\omega > \omega_L$ and substitute Eq. (10.2.11) into Eq. (10.2.9) and use Eq. (10.2.1) to obtain finally

$$F(\omega) = - \frac{\beta \hbar}{1 - e^{-\beta \hbar \omega}} \frac{1}{\pi} \operatorname{Im} \lim_{\delta \rightarrow 0+} \sum_j \frac{2\epsilon \omega^2}{\beta \hbar}$$

$$\begin{aligned}
& \times \frac{1}{N} \frac{\omega_j(k)}{(\omega_j^2(k) - \omega^2)^2} \\
& \times \frac{1}{(\omega^2 + i\delta - \omega_0^2) - \frac{\epsilon}{3N} \sum_{\substack{k_1 \\ j_1}} \frac{\omega_{j_1}^2(k_1)}{(\omega^2 - \omega_{j_1}^2(k_1))^2}}
\end{aligned} \tag{10.2.12}$$

where ω_0^2 is the solution of the equation

$$\begin{aligned}
0 &= 1 + \frac{\lambda\beta\hbar}{2} \frac{1}{3N} \sum_{k,j} \omega_j(k) \alpha_L \left(\frac{k}{j}; i\omega_L = \omega \right) \\
&= 1 - \frac{\epsilon\omega^2}{3N} \sum_{k,j} \frac{1}{\omega^2 - \omega_j^2(k)}
\end{aligned} \tag{10.2.13}$$

We readily identify this equation as the equation for the frequency of the local mode arising from an isolated mass defect. We thus see that $F(\omega)$ has a δ -function singularity at the local mode frequency.

What we must now do is to determine the form of the coefficient $\alpha_L(kj)$ for an anharmonic crystal containing a single mass defect.

This can be done directly from Eq. (10.2.10) if for $\alpha_L(kj)$ and

$\beta_L(kj)$ we substitute their values in an anharmonic crystal. Actually, our task is simpler than this since we have a fairly limited objective. We want to find how the δ -function singularity in $F(\omega)$ at $\omega = \omega_0$ is modified by anharmonic effects. This means that we only have to study Eq. (10.2.13) again, but with $\alpha_L(kj)$ replaced by its form for an anharmonic crystal.

The expression for $\alpha_L(kj)$ in an anharmonic crystal has

recently been determined in connection with another problem and is^{52, 68}

$$\alpha_l \left(\begin{matrix} k \\ j \end{matrix} \right) = \frac{2\omega_j(k)}{\beta\hbar} \frac{1}{\omega_j^2(k) + \omega_l^2 - \frac{2\omega_j(k)}{\beta\hbar} G_l \left(\begin{matrix} k \\ j \end{matrix} \right)}$$

(10.2.14)

where

$$G_l \left(\begin{matrix} k \\ j \end{matrix} \right) = \frac{\beta\hbar^2}{16N\omega_j(k)} \sum_{\substack{k_1 k_2 \\ j_1 j_2}} \Delta(-k + k_1 + k_2) \\ \times \frac{\left| \Phi \left(\begin{matrix} -k & k_1 & k_2 \\ j & j_1 & j_2 \end{matrix} \right) \right|^2}{\omega_{j_1}(k_1)\omega_{j_2}(k_2)} \left\{ \frac{n_1 + n_2 + 1}{i\omega_l + \omega_1 + \omega_2} - \frac{n_1 + n_2 + 1}{i\omega_l - \omega_1 - \omega_2} \right. \\ \left. + \frac{n_1 - n_2}{i\omega_l - \omega_1 + \omega_2} - \frac{n_1 - n_2}{i\omega_l + \omega_1 - \omega_2} \right\}$$

(10.2.15)

and where for convenience we have put $\omega_{j_1}(k_1) = \omega_1$, etc. Only

the cubic anharmonic contribution to the proper self-energy function $G_l(k, j)$ has been retained, since to this order in the anharmonic force constants the quartic terms do not lead to any lifetime effects.⁶⁸ The function we must study is

$$D(\nu) = 1 + \frac{\lambda}{3N} \sum_{\mathbf{k}} \frac{\omega_{\mathbf{j}}^2(\mathbf{k})}{\omega_{\mathbf{j}}^2(\mathbf{k}) - \nu^2 - \frac{2\omega_{\mathbf{j}}(\mathbf{k})}{\beta\hbar} G_{\ell} \left(\frac{\mathbf{k}}{\mathbf{j}}; i\omega_{\ell} = \nu \right)} \quad (10.2.16)$$

when ν is replaced by $-\omega - i\delta$ and the limit $\delta \rightarrow 0 + i\delta$ is taken.

The function $-(1/\beta\hbar) G_{\ell} \left(\frac{\mathbf{k}}{\mathbf{j}}; i\omega_{\ell} \rightarrow \nu \rightarrow -\omega - i\delta \right)$ becomes

$$-\frac{1}{\beta\hbar} G_{\ell} \left(\frac{\mathbf{k}}{\mathbf{j}}; i\omega_{\ell} = -\omega - i\delta \right) = \Delta_{\mathbf{j}}(\mathbf{k}; \omega) + i\Gamma_{\mathbf{j}}(\mathbf{k}; \omega) \quad (10.2.17)$$

where

$$\Delta_{\mathbf{j}}(\mathbf{k}; \omega) = \frac{\hbar}{16N\omega_{\mathbf{j}}(\mathbf{k})} \sum_{\substack{\mathbf{k}_1 \mathbf{k}_2 \\ \mathbf{j}_1 \mathbf{j}_2}} \Delta(-\mathbf{k} + \mathbf{k}_1 + \mathbf{k}_2) \frac{\left| \frac{\hbar}{2} \frac{-\mathbf{k} \mathbf{k}_1 \mathbf{k}_2}{\mathbf{j} \mathbf{j}_1 \mathbf{j}_2} \right|^2}{\omega_{\mathbf{j}_1}(\mathbf{k}_1) \omega_{\mathbf{j}_2}(\mathbf{k}_2)}$$

$$\times \left\{ (n_1 + n_2 + 1) \left[\frac{-1}{(\omega + \omega_1 + \omega_2)_p} + \frac{1}{(\omega - \omega_1 - \omega_2)_p} \right] \right.$$

$$\left. + (n_1 - n_2) \left[\frac{-1}{(\omega - \omega_1 + \omega_2)_p} + \frac{1}{(\omega + \omega_1 - \omega_2)_p} \right] \right\}$$

(10.2.18a)

$$\begin{aligned}
\Gamma_j(\mathbf{k}; \omega) = & \frac{\pi \hbar}{16N\omega_j(\mathbf{k})} \sum_{\substack{\mathbf{k}_1 \mathbf{k}_2 \\ j_1 j_2}} \Delta(-\mathbf{k} + \mathbf{k}_1 + \mathbf{k}_2) \frac{\left| \begin{smallmatrix} -\mathbf{k} & \mathbf{k}_1 & \mathbf{k}_2 \\ j & j_1 & j_2 \end{smallmatrix} \right|^2}{\omega_{j_1}(\mathbf{k}_1) \omega_{j_2}(\mathbf{k}_2)} \\
& \times \left\{ (n_1 + n_2 + 1) \left[-\delta(\omega + \omega_1 + \omega_2) + \delta(\omega - \omega_1 - \omega_2) \right] \right. \\
& \left. + (n_1 - n_2) \left[-\delta(\omega - \omega_1 + \omega_2) + \delta(\omega + \omega_1 - \omega_2) \right] \right\}
\end{aligned}$$

(10.2.18b)

We thus find that

$$\begin{aligned}
\lim_{\delta \rightarrow 0+} D(\nu = -\omega - i\delta) = & 1 + \frac{\lambda}{3N} \sum_{\mathbf{k} j} \\
& \times \frac{\omega_j^2(\mathbf{k}) \left[\omega_j^2(\mathbf{k}) - \omega^2 + 2\omega_j(\mathbf{k}) \Delta_j(\mathbf{k}; \omega) \right]}{\left[\omega_j^2(\mathbf{k}) - \omega^2 + 2\omega_j(\mathbf{k}) \Delta_j(\mathbf{k}; \omega) \right]^2 + 4\omega_j^2(\mathbf{k}) \Gamma_j^2(\mathbf{k}; \omega)} \\
& + i \frac{\lambda}{3N} \sum_{\mathbf{k} j} \frac{2\omega_j^3(\mathbf{k}) \Gamma_j(\mathbf{k}; \omega)}{\left[\omega_j^2(\mathbf{k}) - \omega^2 + 2\omega_j(\mathbf{k}) \Delta_j(\mathbf{k}; \omega) \right]^2 + 4\omega_j^2(\mathbf{k}) \Gamma_j^2(\mathbf{k}; \omega)}
\end{aligned}$$

(10.2.19)

$$= D_1(\omega^2) + i D_2(\omega^2)$$

(10.2.20)

Let us define a frequency ω_{0A} by the equation

$$D_1(\omega_{0A}^2) = 0 \quad (10.2.21)$$

Then in the neighborhood of this frequency we have

$$\lim_{\delta \rightarrow 0+} D(\nu = -\omega - i\delta) = (\omega^2 - \omega_{0A}^2) D_1'(\omega_{0A}^2) + i D_2(\omega_{0A}^2) \quad (10.2.22)$$

This result, together with Eqs. (10.2.10) and (10.2.9), implies that in the neighborhood of ω_{0A}^2 , $F(\omega)$ has the form

$$\begin{aligned} F(\omega) &= \text{const.} \frac{D_2(\omega_{0A}^2)}{(\omega^2 - \omega_{0A}^2) D_1'(\omega_{0A}^2) + D_2(\omega_{0A}^2)} \\ &= \text{const.} \frac{1}{(\omega - \omega_{0A})^2 + \Gamma^2} \end{aligned} \quad (10.2.23)$$

where

$$\Gamma = \frac{D_2(\omega_{0A}^2)}{2\omega_{0A} D_1'(\omega_{0A}^2)} \quad (10.2.24)$$

The width of this resonance at half-maximum is 2Γ , and we obtain finally that the lifetime τ of the localized mode is given by

$$\frac{1}{\tau} = 2\Gamma = \frac{D_2(\omega_{0A}^2)}{\omega_{0A} D_1'(\omega_{0A}^2)} \quad (10.2.25)$$

As it stands, Eq. (10.2.25) is more complicated than is justified by our retention of only cubic anharmonic terms in the proper self-energy $G_L(k; \omega)$. We are justified in obtaining Γ only to $O(\eta^2)$ since higher-order anharmonic terms which we have neglected will contribute to $O(\eta^4)$ and higher. We, accordingly, make the following approximations in obtaining a simpler expression for Γ . First, we neglect the difference between the frequency ω_0 defined by Eq. (10.2.13) and ω_{0A} defined by Eqs. (10.2.19) and (10.2.21). This difference is $O(\eta^2)$ and would be of interest to know because it gives the temperature dependence of the local mode frequency. However, $D_2(\omega^2)$ is $O(\eta^2)$ already, and the replacement of ω_{0A} by ω_0 in this function leads to an error that is $O(\eta^4)$, and therefore is of higher order in η than we consider here. Second, we neglect the terms in $\Gamma_j(k; \omega)$ and $\Delta_j(k; \omega)$ appearing in the denominator of the integral that defines $D_2(\omega^2)$, since these contribute higher-order corrections to the result toward which we are working. We thus obtain the following approximate expressions for $D_1(\omega^2)$ and $D_2(\omega^2)$:

$$\begin{aligned}
 D_1(\omega^2) &\cong 1 + \frac{\lambda}{3N} \sum_j \frac{\omega_j^2(k)}{\omega_j^2(k) - \omega^2} \\
 &= \frac{1}{1 - \epsilon} \left\{ 1 - \frac{\epsilon \omega^2}{3N} \sum_j \frac{1}{\omega^2 - \omega_j^2(k)} \right\}
 \end{aligned}
 \tag{10.2.26a}$$

$$D_1'(\omega^2) \cong \frac{1}{1 - \epsilon} - \frac{\epsilon}{3N} \sum_j \frac{\omega_j^2(k)}{(\omega^2 - \omega_j^2(k))^2}
 \tag{10.2.26b}$$

$$D_2(\omega^2) \cong \frac{1}{1 - \epsilon} - \frac{2\epsilon}{3N} \sum_j \frac{\omega_j^3(k) \Gamma_j(k; \omega)}{(\omega^2 - \omega_j^2(k))^2}
 \tag{10.2.26c}$$

The phonon lifetime is therefore given to $O(\eta^2)$ by

$$\frac{1}{\tau} = \frac{2}{3N} \sum_{\mathbf{k}, \mathbf{j}} \frac{\omega_{\mathbf{j}}^3(\mathbf{k}) \Gamma_{\mathbf{j}}(\mathbf{k}; \omega_0)}{(\omega_0^2 - \omega_{\mathbf{j}}^2(\mathbf{k}))^2} \times \left\{ \frac{\omega_0}{3N} \sum_{\mathbf{k}, \mathbf{j}} \frac{\omega_{\mathbf{j}}^2(\mathbf{k})}{(\omega_0^2 - \omega_{\mathbf{j}}^2(\mathbf{k}))^2} \right\}^{-1} \quad (10.2.27)$$

We now turn to the approximate evaluation of this expression.

3. Numerical Estimate of the Lifetime of a Localized Mode

Our first task is to evaluate $\Gamma_{\mathbf{j}}(\mathbf{k}; \omega_0)$. Since the lifetime decreases with increasing temperature, as can be seen from the high-temperature form of $\Gamma_{\mathbf{j}}(\mathbf{k}; \omega)$, and since we are interested in seeing how long the lifetime can be, we evaluate τ^{-1} at the absolute zero of temperature. In this limit we see from Eq. (10.2.18b) that $\Gamma_{\mathbf{j}}(\mathbf{k}; \omega_0)$ is given by

$$\Gamma_{\mathbf{j}}(\mathbf{k}; \omega_0) = \frac{\pi \hbar}{16N\omega_{\mathbf{j}}(\mathbf{k})} \sum_{\substack{\mathbf{k}_1, \mathbf{k}_2 \\ \mathbf{j}_1, \mathbf{j}_2}} \Delta(-\mathbf{k} + \mathbf{k}_1 + \mathbf{k}_2) \left| \frac{\begin{pmatrix} -\mathbf{k} & \mathbf{k}_1 & \mathbf{k}_2 \\ \mathbf{j} & \mathbf{j}_1 & \mathbf{j}_2 \end{pmatrix}}{\omega_{\mathbf{j}_1}(\mathbf{k}_1)\omega_{\mathbf{j}_2}(\mathbf{k}_2)} \right|^2 \delta(\omega_0 - \omega_1 - \omega_2) \quad (10.3.1)$$

An exact evaluation of this expression is quite difficult at the present time. We therefore make a number of simplifying assumptions that will

enable us to carry out the calculation analytically and which probably yield a result that is in order of magnitude agreement with the results of an exact calculation.

It is convenient to replace summation over k by integration throughout the Brillouin zone according to

$$\sum_k \rightarrow V \int_{Bz} d^3k \quad (10.3.2)$$

where V is the volume of the crystal and equals Nv_a ; v_a is the volume of a unit cell. The discrete Δ function goes over into a sum of Dirac Δ functions,

$$\Delta(-k + k_1 + k_2) = \frac{1}{V} \sum_{\tau} \delta(-k + k_1 + k_2 + \tau) \quad (10.3.3)$$

where τ is a translation vector of the reciprocal lattice. In what follows we make the approximation of retaining only the $\tau = 0$ term in Eq. (10.3.3). This means that we ignore the contribution from the so-called "Umklapp" processes.

The second approximation we make is to use an approximate form for $\Phi(-k, j; k_1, j_1; k_2, j_2)$ due to Klemens,⁶⁹

$$\left| \Phi \begin{pmatrix} -k & k_1 & k_2 \\ j & j_1 & j_2 \end{pmatrix} \right|^2 \approx 48 \frac{\gamma^2}{Mc^2} \omega_j^2(k) \omega_{j_1}^2(k_1) \omega_{j_2}^2(k_2) \quad (10.3.4)$$

where γ is the Grüneisen constant, and c is an average speed of sound.

The final approximation we make is to use the Debye approximation for the frequencies

$$\omega_j(k) = 2\pi c k \quad (10.3.5)$$

where the sound speed c is assumed to be the same for all branches and

to replace the Brillouin zone by a sphere of equal volume, so that its radius k_D is given by

$$\frac{4}{3} \pi k_D^3 = \frac{1}{v_a} \quad (10.3.6)$$

With these approximations we obtain

$$\begin{aligned} \Gamma_j(k; \omega_0) = & 27\pi \gamma_a^2 v_a \left(\frac{\hbar \omega_j(k)}{Mc^2} \right) 2\pi c \int d^3 k_1 \int d^3 k_2 k_1 k_2 \\ & \times \delta(-k + k_1 + k_2) \delta(k_0 - k_1 - k_2) \end{aligned} \quad (10.3.7)$$

where

$$k_0 = \frac{\omega_0}{2\pi c} > k_D \quad (10.3.8)$$

Denoting the double integral in Eq. (10.3.7) by I , we have

$$\begin{aligned} I = & \int_{\substack{k_1 < k_D \\ |k - k_1| < k_D}} d^3 k_1 k_1 |k - k_1| \delta(k_0 - k_1 - |k - k_1|) \\ = & 2\pi \int_0^{k_D} dk_1 k_1^3 \int_{-1}^1 dx \sqrt{k^2 + k_1^2 - 2k k_1 x} \\ & \times \delta\left(k_0 - k_1 - \sqrt{k^2 + k_1^2 - 2k k_1 x}\right) \end{aligned}$$

$$= 2\pi \int_0^{k_D} dk_1 k_1^3 (k_0 - k_1) \int_{-1}^1 dx \frac{\delta(x - x_0)}{\left(\frac{k k_1}{k_0 - k_1} \right)} \sqrt{k^2 + k_1^2 - 2k k_1 x_0} < k_D$$

(10.3.9)

where x_0 is the root of

$$\sqrt{k^2 + k_1^2 - 2k k_1 x_0} = k_0 - k_1$$

(10.3.10)

and must lie in the interval $(-1, 1)$. The integral over x in Eq. (10.3.9) is readily carried out, and we are left with

$$I = \frac{2\pi}{k} \int dk_1 k_1^2 (k_0 - k_1)^2$$

(10.3.11)

with the following restrictions on k_1 :

$$0 < k_1 < k_D$$

(10.3.12a)

$$k_0 - k_1 < k_D$$

(10.3.12b)

$$(k + k_1)^2 > (k_0 - k_1)^2 > (k - k_1)^2$$

(10.3.12c)

The first of these restrictions is just the usual one on the magnitude of k_1 . The second follows from the restriction $|k - k_1| < k_D$, and the third follows from the condition $-1 < x_0 < 1$. We now consider the two inequalities expressed by Eq. (10.3.12c). After some manipulation they yield the conditions:

$$k_1 > \frac{1}{2} (k_0 - k) \quad (10.3.13a)$$

$$k_1 > \frac{1}{2} (k_0 + k) \quad (10.3.13b)$$

Clearly, the second condition is more restrictive than the first. Equation (10.3.12b) tells us

$$k_1 > k_0 - k_D$$

However, we readily find that

$$\frac{1}{2} (k_0 + k) > k_0 - k_D$$

since $k_0 < 2k_D$, so that the restrictions on k_1 become

$$\frac{1}{2} (k_0 + k) < k_1 < k_D \quad (10.3.14)$$

provided that

$$\frac{1}{2} (k_0 + k) < k_D \quad \text{or} \quad k < 2k_D - k_0 \quad (10.3.15)$$

The integral vanishes otherwise.

We finally obtain the result

$$I = \frac{2\pi}{k} \int_{\frac{1}{2}(k_0+k)}^{k_D} dk_1 \left(k_0^2 k_1^2 - 2k_0 k_1^3 + k_1^4 \right)$$

$$= \frac{2\pi}{\omega} \frac{1}{(2\pi c)^4} \left\{ \frac{1}{3} \omega_0^2 \left[\omega_D^3 - \left(\frac{\omega_0 + \omega}{2} \right)^3 \right] \right\}$$

$$- \frac{1}{2} \omega_0 \left[\omega_D^4 - \left(\frac{\omega_0 + \omega}{2} \right)^4 \right] + \frac{1}{5} \left[\omega_D^5 - \left(\frac{\omega_0 + \omega}{2} \right)^5 \right] \Bigg\}$$

$$0 < \omega < 2\omega_D - \omega_0$$

$$= 0 \quad \text{otherwise}$$

(10.3.16)

The expression for $\Gamma_j(k; \omega_0)$ is found from Eqs. (10.3.7) and (10.3.16) to be

$$\begin{aligned} \Gamma_j(k; \omega_0) &= \frac{81\pi\gamma^2}{2} \left(\frac{\hbar\omega_D}{Mc^2} \right) \frac{1}{\omega_D^4} \left\{ \frac{\omega_0^2}{3} \right. \\ &\times \left[\omega_D^3 - \left(\frac{\omega_0 + \omega}{2} \right)^3 \right] - \frac{\omega_0}{2} \left[\omega_D^4 - \left(\frac{\omega_0 + \omega}{2} \right)^4 \right] \\ &\left. + \frac{1}{5} \left[\omega_D^5 - \left(\frac{\omega_0 + \omega}{2} \right)^5 \right] \right\} \quad 0 < \omega < 2\omega_D - \omega_0 \\ &= 0 \quad \text{otherwise} \end{aligned}$$

(10.3.17)

In writing Eq. (10.3.17) we have used Eq. (10.3.6).

The numerator in Eq. (10.2.27) can now be written

$$243 \pi \gamma^2 \left(\frac{\hbar\omega_D}{Mc^2} \right) I_1(x_0)$$

(10.3.18)

where

$$I_1(x_0) = \int_0^{2-x_0} dx \frac{x^5}{(x_0^2 - x^2)^2} \left\{ \frac{x_0^2}{3} \left[1 - \left(\frac{x_0 + x}{2} \right)^3 \right] - \frac{x_0}{2} \left[1 - \left(\frac{x_0 + x}{2} \right)^4 \right] + \frac{1}{5} \left[1 - \left(\frac{x_0 + x}{2} \right)^5 \right] \right\} \quad (10.3.19)$$

We have made the definition $w_0 = \frac{w_0}{D} x_0$.

We introduce the parameter

$$1 > y = \frac{2 - x_0}{x_0} > 0 \quad (10.3.20)$$

and rewrite $I_1(x_0)$ as

$$I_1(x_0) = \frac{10 x_0^4 - 15 x_0^3 + 6 x_0^2}{30} F_1(y) - \frac{x_0^7}{480} F_2(y) \quad (10.3.21)$$

where

$$F_1(y) = \int_0^y \frac{z^5}{(1-z^2)^2} dz = \frac{1}{2} y^2 \left(\frac{2-y^2}{1-y^2} \right) + \ln(1-y^2) \quad (10.3.22)$$

$$\begin{aligned} F_2(y) &= \int_0^y \frac{z^5}{(1-z^2)^2} (8 + 15z - 10z^3 + 3z^5) dz \\ &= \frac{4y}{1-y} + 16 \ln(1-y) + 12y + 4y^2 + \frac{4}{3} y^3 - \frac{4}{5} y^5 + \frac{3}{7} y^7 \end{aligned} \quad (10.3.23)$$

Table X. 1 gives the function $I_1(x_0)$ in brief.

The integral in the denominator of Eq. (10. 2. 27) is given by

$$\begin{aligned} \frac{1}{\omega_D} I_2(x_0) &= \frac{3x_0}{\omega_D} \int_0^1 \frac{x^4}{(x_0^2 - x^2)^2} dx \\ &= \frac{1}{\omega_D} \left\{ 3x_0 - \frac{9}{4}x_0^2 \ln \frac{x_0 + 1}{x_0 - 1} + \frac{3}{2} \frac{x_0^3}{x_0^2 - 1} \right\} \end{aligned} \quad (10. 3. 24)$$

The function $I_2(x_0)$ is also tabulated in Table X. 1.

Combining Eqs. (10. 3. 18) and (10. 3. 24) we obtain finally that the life-time of the localized mode is given by

$$\frac{1}{\tau} = 243\pi\omega_D \gamma^2 \left(\frac{\hbar\omega_D}{Mc^2} \right) \frac{I_1(x_0)}{I_2(x_0)} \quad (10. 3. 25)$$

See Table X. 1 for the tabulation of the quotient $I_1(x_0)/I_2(x_0)$.

To complete our analysis we must relate the local mode frequency $x_0 = \omega_0/\omega_D$ to the mass defect parameter ϵ . x_0 is given by the solution of the equation

$$\ln \frac{x_0 + 1}{x_0 - 1} = \frac{2(1 + 3\epsilon x_0^2)}{3\epsilon x_0^3} \quad (10. 3. 26)$$

Values of ϵ as a function of x_0 are tabulated in Table X. 1.

In a typical case, $\gamma \approx 2$, and $(\hbar\omega_D/Mc^2) \approx 1/100^9$. These values imply that

$$\frac{1}{\tau} \approx \omega_D (30.5) \frac{I_1(x_0)}{I_2(x_0)} \quad (10. 3. 27)$$

Table X.1. Functions Required for the Calculation of the Lifetime of a Localized Vibration Mode

x_0	y	$F_1(y)$	$F_2(y)$	$I_1(x_0)$	$I_2(x_0)$	$\frac{I_1(x_0)}{I_2(x_0)}$	ϵ	$\frac{M'}{M}$
1.1	0.818182	2.403×10^{-1}	3.762	2.340×10^{-4}	4.518	5.179×10^{-5}	0.408	0.592
1.2	0.666667	3.444×10^{-2}	5.148×10^{-1}	1.242×10^{-4}	1.722	7.212×10^{-5}	0.526	0.474
1.3	0.538462	6.730×10^{-3}	9.451×10^{-2}	5.362×10^{-5}	0.931	5.759×10^{-5}	0.610	0.390
1.4	0.428571	1.396×10^{-3}	1.823×10^{-2}	1.927×10^{-5}	0.586	3.288×10^{-5}	0.667	0.333
1.5	0.333333	2.725×10^{-4}	3.289×10^{-3}	5.552×10^{-6}	0.402	1.381×10^{-5}	0.714	0.286
1.6	0.250000	4.481×10^{-5}	4.984×10^{-4}	1.189×10^{-6}	0.292	4.071×10^{-6}	0.752	0.248
1.7	0.176471	5.278×10^{-6}	5.403×10^{-5}	1.607×10^{-7}	0.221	7.271×10^{-7}	0.781	0.219
1.8	0.111111	3.195×10^{-7}	3.010×10^{-6}	9.470×10^{-9}	0.172	5.506×10^{-8}	0.806	0.194
1.9	0.052632	3.544×10^{-9}	3.086×10^{-8}	7.432×10^{-11}	0.138	5.386×10^{-10}	0.826	0.174
2.0	0	0	0	0	0.113	0	0.847	0.153

From this result and Table X, 1 we see that for $x_0 \cong 1.5$,

$$\frac{1}{\tau} \cong \omega_D (0.4 \times 10^{-3}) \quad (10.3.28)$$

For $x_0 = 1.2$, our result for τ^{-1} agrees with the result of Klemens to within a factor of about 2, but seems to decrease more rapidly with increasing x_0 than does his result.

Acknowledgments

This report describes the results of work carried out at various times primarily by the author in collaboration with several colleagues, whose contributions he should like to acknowledge here. The material in Sec. III is based on work by J. Mahanty, G. H. Weiss, and myself. Section IV contains results obtained together with P. A. Flinn and S. Ruby, and some of the work in Sec. IX is an outgrowth of work carried out with P. A. Flinn. Section V is based on as yet unpublished work done in collaboration with G. H. Weiss and C. T. Chen. In particular, I should like to acknowledge the collaboration of J. M. Radcliffe in the work described in Secs. VII through X. I am grateful to C. Domb, C. Isenberg, R. Davies, and W. Visscher for sending me accounts of their work prior to its publication.

References

1. An extensive bibliography of the work of Lifshitz and his collaborators is given in I. M. Lifshitz, *Nuovo Cimento, Suppl.*, [10], 3, 716 (1956).
2. E. W. Montroll and R. B. Potts, *Phys. Rev.*, 100, 525 (1955); P. Mazur, E. W. Montroll, and R. B. Potts, *J. Wash. Acad. Sci.*, 46, 2 (1956); E. W. Montroll and R. B. Potts, *Phys. Rev.*, 102, 72 (1956); E. W. Montroll, A. A. Maradudin, and G. H. Weiss, "Proceedings of the Stevens Institute Conference on the Many-Body Problem," Interscience (to be published); A. A. Maradudin, P. Mazur, E. W. Montroll, and G. H. Weiss, *Revs. Mod. Phys.*, 30, 175 (1958).
3. O. Litzman, *Czech. J. Phys.*, 7, 410 (1957); *ibid.*, 7, 690 (1957); *ibid.*, 8, 521 (1958); *ibid.*, 8, 633 (1958); *ibid.*, 9, 692 (1959).
4. J. Hori and T. Asahi, *Progr. Theoret. Phys. (Kyoto)*, 17, 523 (1957); J. Hori, *ibid.*, 18, 367 (1957); J. Hori, *J. Phys. Soc. Japan*, 16, 23 (1961).

5. E. Teramoto, *Progr. Theoret. Phys.*, 24, 1296 (1960); E. Teramoto and S. Takeno, *ibid.*, 24, 1349 (1960); *ibid.*, 25, 102 (1961).
6. P. Dean, *Proc. Phys. Soc. (London)*, 73, 413 (1959); P. Dean, *Proc. Roy. Soc. (London)*, A254, 507 (1960); P. Dean, *ibid.*, A260, 264 (1961).
7. P. Dean and J. L. Martin, *Proc. Roy. Soc. (London)*, A259, 409 (1960).
8. P. Wojtowicz and J. G. Kirkwood, *J. Chem. Phys.*, 33, 1299 (1960).
9. P. G. Klemens, *Phys. Rev.*, 122, 443 (1961).
10. P. G. Klemens, *Proc. Phys. Soc. (London)*, A68, 1113 (1955); P. Carruthers, *Phys. Rev.*, 126, 1448 (1962).
11. I. M. Lifshitz, *Zhur. Eksperim. i Teoret. Fiz.*, 18, 293 (1948).
12. G. F. Koster, *Phys. Rev.*, 95, 1436 (1954).
13. A. A. Maradudin, G. H. Weiss, and D. W. Jepsen, *J. Math. Phys.*, 2, 349 (1960); J. Hori, *J. Math. Phys.*, 3, 382 (1962).
14. E. P. Adams, "Smithsonian Mathematical Formulae and Tables of Elliptic Functions," p. 84, The Smithsonian Institution, Washington, D. C., 1957.
15. Lord Rayleigh, "The Theory of Sound," Vol. I, Dover Publications, New York, 1945.
16. A survey of such methods is given in A. A. Maradudin, E. W. Montroll, and G. H. Weiss, "The Theory of Lattice Dynamics in the Harmonic Approximation," Suppl. Vol. No. 3, "Solid State Physics," (F. Seitz and D. Turnbull, eds.), Academic Press, New York, 1963.
17. G. F. Nardelli and N. Tettamanzi, *Phys. Rev.*, 126, 1283 (1962).
18. J. Mahanty, A. A. Maradudin, and G. H. Weiss, *Progr. Theoret. Phys. (Kyoto)*, 24, 648 (1960).
19. R. A. Sack, A. A. Maradudin, and G. H. Weiss, *Phys. Rev.*, 124, 717 (1961); C. Domb and C. Isenberg, *Proc. Phys. Soc. (London)*, 79, 659 (1962).
20. K. Yamahuzi and T. Tanaka, *Progr. Theoret. Phys. (Kyoto)*, 20, 327 (1950).
21. M. Lax, *Phys. Rev.*, 94, 1391 (1954).
22. J. C. Phillips, *Phys. Rev.*, 104, 1263 (1956).
23. C. Domb, A. A. Maradudin, E. W. Montroll, and G. H. Weiss, *Phys. Rev.*, 115, 24 (1959).
24. C. Domb and C. Isenberg (private communication).
25. See, for example, A. A. Maradudin, P. A. Flinn, and R. A. Coldwell-Horsfall, *Ann. Phys. (N.Y.)*, 15, 360 (1961).
26. A. A. Maradudin, P. A. Flinn, and S. Ruby, *Phys. Rev.*, 126, 9 (1962).
27. G. K. Horton and H. Schiff, *Phys. Rev.*, 104, 32 (1956).
28. P. Marcus (private communication).

29. C. Domb and L. Salter, *Phil. Mag.*, 43, 1083 (1952).
30. B. D. Josephson, *Phys. Rev. Letters*, 4, 341 (1960). An earlier derivation of the magnitude of this shift was given by R. V. Pound and G. A. Rebka, Jr., *Phys. Rev. Letters*, 4, 274 (1960).
31. R. Mössbauer, *Z. Physik*, 151, 124 (1958); *Naturwissenschaften*, 45, 538 (1958); *Z. Naturforsch.*, 14a, 211 (1959).
- 31a. L. R. Walker, G. K. Wertheim, and V. Jaccarino, *Phys. Rev. Letters*, 6, 98 (1961).
32. R. W. Hill and P. L. Smith, *Phil. Mag.*, (7) 44, 636 (1953).
33. I. Prigogine, R. Bingen, and J. Jeener, *Physica*, 20, 388 (1954).
34. G. V. Chester, *Phys. Rev.*, 100, 446 (1955).
35. I. M. Lifshitz and G. I. Stepanova, *Soviet Phys. JETP (English Transl.)*, 3, 656 (1956).
36. G. H. Weiss and A. A. Maradudin, *Phys. Chem. Solids*, 7, 327 (1958).
37. C. T. Chen, Ph. D. Thesis, University of Maryland, 1962 (unpublished).
38. J. S. Langer, *J. Math. Phys.*, 2, 584 (1961).
39. F. J. Dyson, *Phys. Rev.*, 92, 1331 (1953).
40. R. Bellman, *Phys. Rev.*, 101, 19 (1956).
41. R. Englman, *Nuovo Cimento*, 10, 615 (1958).
42. H. Schmidt, *Phys. Rev.*, 105, 425 (1957).
43. R. J. Rubin and B. A. Fleishman, *Phys. Rev.*, 99, 656 (1955); (private communication).
44. C. Domb, A. A. Maradudin, E. W. Montroll, and G. H. Weiss, *Phys. Rev.*, 115, 18 (1959).
45. J. Hori, *Progr. Theoret. Phys. (Kyoto)*, 23, 475 (1960).
46. W. C. Overton, Jr., and E. Dent, U. S. Naval Research Lab. Rep. 5252, Washington, D. C. (unpublished).
47. R. Davies (private communication).
48. J. L. Martin, *Proc. Roy. Soc. (London)*, A260, 139 (1961).
49. M. D. Bacon and P. Dean, *Nature*, 194, No. 4828, p. 541 (1962).
50. H. B. Rosenstock and R. E. McGill, *J. Math. Phys.*, 3, 200 (1962).
51. P. A. Flinn and A. A. Maradudin (unpublished work).
52. A. A. Maradudin, "The Anharmonic Solid" Vol. II, Lecture Notes, Physics Department, Carnegie Institute of Technology, Spring Semester, 1962 (unpublished).
53. C. Bloch and C. DeDominicis, *Nuclear Phys.*, 7, 459 (1958).
54. R. Kubo, *J. Phys. Soc. Japan*, 12, 570 (1957).
- 54a. G. Baym and N. D. Mermin, *J. Math. Phys.*, 2, 232 (1961).
55. M. Born and K. Huang, "Dynamical Theory of Crystal Lattices," p. 267, Oxford Univ. Press, London, 1954.
56. R. F. Wallis and A. A. Maradudin, Proceedings of the 1962 International Conference on Semiconductor Physics, Exeter (to be published).
57. G. Schaefer, *Phys. Chem. Solids*, 12, 233 (1960).

58. W. E. Lamb, Phys. Rev., 55, 190 (1939). Strictly speaking, Lamb's expression, which was derived for the case of resonance absorption of neutrons of energy E by a nucleus in a crystal, differs from Eq. (9.1.1) by a constant factor.
59. H. F. Baker, Proc. London Math. Soc., 3, 24 (1905); F. Hausdorff, Ber. Sächsischen Akad. Wiss. (Math. Phys. Kl.) Leipzig, 58, 19 (1906).
60. See, for example, L. Van Hove, Phys. Rev., 95, 249 (1954); B. Kaufman and H. J. Lipkin, Ann. Phys. (N.Y.), 18, 294 (1962).
61. P. Debye, Ann. Physik, 43, 49 (1914); I. Waller, Dissertation, Uppsala, Sweden (1925).
62. K. S. Singwi and A. Sjölander, Phys. Rev., 120, 1093 (1960).
63. A. A. Maradudin and P. A. Flinn, Phys. Rev., 126, 2059 (1962).
64. W. Visscher, Ann. Phys. (N.Y.), 9, 194 (1960).
65. S. V. Maleev, Soviet Phys. JETP (English Transl.), 12, 617 (1961); W. Visscher, in "Proceedings of the Second Mössbauer Conference, Paris, 1961," Wiley, New York, 1962; B. Mozer and G. H. Vineyard, Bull. Am. Phys. Soc., [II] 6, 135 (1961).
66. R. Brout and W. Visscher, Phys. Rev. Letters, 9, 54 (1962).
67. L. Van Hove, Tech. Rept. No. 11, Solid State and Molecular Theory Group, M.I.T., March 15, 1959. This report is reprinted in L. Van Hove, N. M. Hugenholtz, and L. P. Howland, "Quantum Theory of Many-Particle Systems" (lecture notes and selected reprints), W. A. Benjamin, New York, 1961.
68. A. A. Maradudin and A. E. Fein, Phys. Rev., 128, 2589 (1962).
69. P. G. Klemens, in "Solid State Physics," Vol. 7, p. 1 (F. Seitz and D. Turnbull, eds.), Academic Press, New York, 1958.

GREEN'S FUNCTIONS IN MANY-BODY PROBLEMS

V. Ambegaokar

Cornell University

CONTENTS

1. Introduction	323
The Landau Theory	324
Effective Mass Sum Rule	328
The Friedel Sum Rule	328
Transport Properties	330
2. Formal Properties of Green's Functions	333
Theory of the Single-Particle Green's Function	336
Equations of Motion; Self-Energy Operator	345
Perturbation Theory	349
3. Equilibrium Properties	358
The Friedel Sum Rule	361
Thermodynamics	363
4. Transport Properties	371
An Approximate Calculation of Residual Resistivity	377
The Reduced Graph Expansion	400
Vertex Function and Ward Identity	412
Thermal Conductivity and Concluding Remarks	423
Appendix A: Low-Temperature Form of $\Gamma_{\mathbf{k}}(\omega)$	426
Appendix B: Luttinger's Theorem	428
Appendix C: Optical Sum Rule	429
Appendix D: Effective Mass Sum Rule	432
References	437

1. INTRODUCTION

These lectures are intended to serve two purposes: to give an introduction to some methods that are being increasingly used to study systems of many particles; and to show how these methods have been applied to a particular system. I shall not explicitly distinguish between these aims, and, in the main, shall tailor the formal development to the particular applications. However, the generality of the methods should be apparent, and an understanding of what follows should simplify the study of papers in which similar methods are applied to other problems.

The methods I have in mind are adapted from the methods of quantum-field theory. Field theories, of course, are designed to treat systems with many degrees of freedom. The methods we shall be using are based on those developed to handle the profound degrees of freedom associated with the vacuum that manifest themselves at relativistic energies. The degrees of freedom that will interest us are less subtle. They are simply the coordinates of the many particles in the system. As a result, the divergence difficulties of relativistic quantum field theory will not beset us. On the other hand, we shall need to generalize slightly the usual field formulation of relativistic quantum theory for the following reason. We shall wish to consider systems that in general have extremely dense spectra of energy levels and we shall only want to calculate gross properties that are easily susceptible to measurement. We shall not, therefore, be interested in the detailed structure of quantum levels, but only in the behavior of systems statistically distributed over these levels. The generalization of the field theoretical formalism necessary for our purposes has then to do with the taking of statistical averages.

The particular system I shall spend my time discussing is the so-called normal Fermi system. By a normal Fermi system is meant these days a system that behaves more or less like a noninteracting system of fermions. The best known empirical example of such a system is the system of electrons in a nonsuperconducting metal. Another is liquid helium-3 at low temperatures. In fact, until recently it has been something of a puzzle to understand why metallic properties could be so satisfactorily explained by the independent particle models customary in conventional solid state physics. The modern understanding of this puzzle, which is largely due to Landau, ⁽¹⁾ is as follows: The ground state of an interacting many-fermion system will, in general, be a complicated state in which the motions of the different particles are highly correlated. Under some circumstances, however, the low-lying excited states may have a simple structure. In particular, it may be possible to put these excited states, whether they are true eigenstates or long-lived unstable states, into a one-to-one correspondence with the low-lying excitations of the ideal Fermi gas. When this is so, one expects a marked similarity between the interacting and noninteracting Fermi systems for all properties whose measurement only weakly excites the system. Conversely, because metals

do exist, one is led to infer that the energy spectrum of the interacting electrons in a metal is of the kind described above. The low-lying excitations of the interacting system are usually called "quasi particles."

Landau made these ideas the basis of a remarkable intuitive theory, which he called "the theory of the Fermi liquid." I shall give a short description of this theory in the rest of Sec. 1 and use it to calculate some simple equilibrium and transport properties. Thereafter I shall begin to develop the microscopic theory of Green's functions, using essentially the form given by Martin and Schwinger.⁽²⁾ The simple results of the Landau theory will then be derived from the microscopic theory. For the equilibrium properties I shall follow the work of Luttinger,⁽³⁾ and for the transport properties that of Langer.⁽⁴⁾ I should state in advance that, although the microscopic approach is more complete, and in principle applicable when the Landau theory is not, it is no more mathematically rigorous. However, even where we simply duplicate the results of the Landau theory the microscopic approach is instructive. For the Landau theory, as we shall see, proceeds in great leaps of physical insight that are difficult for ordinary mortals to follow. No such difficulties will occur in the more formal approach, and a clear picture of the structure underlying Landau's arguments will emerge as we proceed.

The Landau Theory

The Landau theory begins by postulating a single-particle dispersion curve E_r , where r is a complete set of one-particle quantum numbers consistent with the symmetry of the system. For certain values of r corresponding to excitations in the neighborhood of the ground state, E_r is supposed to be the energy of a quasi-particle excitation. Associated with E_r is a distribution of function f_r^- which is the probability of occupation of the state r . Although no physical meaning is given in the theory to E_r except in the region of r mentioned, it is supposed that the quantum numbers r may be used to enumerate the states of the interacting system. The number of particles in the system, for example, is written

$$N = \sum_r f_r^- \quad (1)$$

The internal energy of the ground state of the systems is not supposed to be expressible in terms of f_r^- and E_r . However for weak excitations, the difference in internal energy of the ground state and the excited states is written

$$\delta E = \sum_r E_r \delta f_r^- \quad (2)$$

The fermion character of the system is introduced by an assumption for the entropy S . It is argued that since the label r serves to enumerate the states, the entropy, which arises solely from the counting of quantum levels, has the same form as that for an ideal gas of fermions, namely,

$$S = -k_B \sum_r \left[(f_r^-) \ln f_r^- + (f_r^+) \ln f_r^+ \right] \quad (3)$$

Here k_B is Boltzmann's constant and

$$f_r^+ = 1 - f_r^- \quad (4)$$

The form of f_r^- in equilibrium is obtained from the variational principle

$$\delta S - \frac{1}{T} \delta E + \frac{\mu}{T} \delta N = 0 \quad (5)$$

where T is the temperature and μ the chemical potential. From (5) one obtains in the usual way

$$f_r^-(T, \mu) = \left(e^{\beta(E_r - \mu)} + 1 \right)^{-1} = \frac{1}{k_B T} \quad (6)$$

From (2) we see that, in general,

$$E_r = \frac{\partial E}{\partial f_r^-} \quad (7)$$

Thus the quasi-particle energy depends on the distribution function and vice versa. In particular, in equilibrium at $T = 0$ we have from (6)

$$f_r^-(0) = \theta(\mu - E_r^0) \quad (8)$$

where E_r^0 is the quasi-particle dispersion curve at $T = 0$ and

$$\theta(x) = \begin{cases} 1 & x > 0 \\ 0 & x < 0 \end{cases} \quad (9)$$

The Landau theory only envisages small variations of f_r^- from (8). For such variations one can expand (7):

$$E_r = E_r^0 + \sum_{r'} \epsilon_{rr'} \delta f_{r'}^- \quad (10)$$

Here

$$\xi_{rr'} = \frac{\delta^2 E}{\delta f_r^- \delta f_{r'}^-} \bigg|_{f^- = f^-(0)} \quad (11)$$

The quantities E_r^0 and $\xi_{rr'}$ are the basic parameters of the Landau theory. They are presumed known. The aim of the theory is to express measurable properties of the system in terms of these parameters and those describing the scattering of quasi particles, and thereby to obtain relations between experimental quantities. Let us see how this may be done in some simple cases. Consider first the specific heat. We have

$$\begin{aligned} C_V &= \left(\frac{\partial E}{\partial T} \right)_{N, V} = \left(\frac{\partial E}{\partial T} \right)_{\mu, V} - \left(\frac{\partial E}{\partial N} \right)_{T, V} \left(\frac{\partial N}{\partial T} \right)_{\mu, V} \\ &= \left(\frac{\partial E}{\partial T} \right)_{\mu, V} - \mu \left(\frac{\partial N}{\partial T} \right)_{\mu, V} \end{aligned} \quad (12)$$

The two partial derivatives in the last line may be calculated from (2) and (1), respectively. For the part of the specific heat linear in T , it is easy to see that one can neglect the first temperature correction (proportional to T^2) in the quasi-particle dispersion curve. The calculation then becomes identical to that for the ideal Fermi gas and one obtains the usual result:

$$C = \frac{\pi^2 k_B^2 T}{3} \sum_r \delta(\mu^0 - E_r^0) \quad (13)$$

Here μ^0 is the chemical potential at $T = 0$. For an homogeneous system the quantum label r becomes the momentum k and spin σ , and the result (13) may be written in terms of the Fermi momentum k_F and the effective mass of a quasi particle at the Fermi surface. These quantities are defined according to

$$\mu^0 = E_{k_F}^0 \quad (14)$$

$$\left. \frac{\partial E}{\partial k} \right|_{k=k_F} = \frac{k_F}{m^*} \quad (15)$$

Then in a straightforward way one has

$$C = \frac{k_B^2}{3} m^* k_F T \quad (16)$$

Another thermodynamic quantity that can easily be calculated is the compressibility. One has[†]

$$K = -\frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_{T, N} = \frac{V}{N^2} \left(\frac{\partial N}{\partial \mu} \right)_{V, T} \quad (17)$$

From (1) and (8) we then have at $T = 0$,

$$\begin{aligned} K &= \frac{V}{N^2} \sum_r \frac{\partial}{\partial \mu^0} \theta(\mu^0 - E_r^0) \\ &= \frac{V}{N^2} \sum_r \left(1 - \frac{\partial E_r^0}{\partial \mu^0} \right) \delta(\mu^0 - E_r^0) \end{aligned} \quad (18)$$

For a free electron gas the second term in the bracket is zero and one has a simple proportionality between the specific heat and the compressibility. This simple relation does not survive in the Landau theory. It is replaced by a more complicated relation involving the interaction function Φ . One has for the translationally invariant case from (10)

$$\left(\frac{\partial E_k}{\partial \mu} \right)_{k=k_F} = \frac{2V}{(2\pi)^3} \int k'^2 dk' d\Omega \Phi(k_F, k') \frac{\partial f_{k'}^-}{\partial \mu}, \quad (19)$$

from which one easily obtains, by writing $\frac{\partial f_{k'}^-}{\partial \mu}$ as indicated in last line of

Eq. (18),

$$K = \left(\frac{V}{N} \right)^2 \frac{k_F}{\pi^2} \left(\frac{1}{m^*} + \frac{2V}{(2\pi)^3} k_F \int d\Omega \Phi(k_F, \theta) \right)^{-1} \quad (20)$$

A final remark. For a free electron gas the spin paramagnetism is also simply proportional to the specific heat. This simple relation is also lost in the Landau theory

[†] The last identity follows from the fact that

$$\mu = \mu \left(\frac{N}{V}, T \right)$$

so that

$$\begin{aligned} \left(\frac{\partial \mu}{\partial N} \right)_{V, T} &= -\frac{V}{N} \left(\frac{\partial \mu}{\partial V} \right)_{N, T} = -\frac{V}{N^2} \left(\frac{\partial (N\mu)}{\partial V} \right)_{N, T} \\ &= -\frac{V^2}{N^2} \left(\frac{\partial P}{\partial V} \right)_{N, T} \end{aligned}$$

Effective Mass Sum Rule

Landau also showed that for a translationally invariant system there is a definite relation between the effective mass m^* and the quantity Φ of Eqs. (10) and (11). He argued that because of Galilean invariance the momentum that arrives at a unit volume must equal the density of mass flow. Assuming, as is natural, that $\partial E_k / \partial k$ is the expression for the velocity of a quasi-particle, one obtains

$$\sum_{\underline{k}} \underline{k} f_{\underline{k}}^- = \sum_{\underline{k}} m \frac{\partial E_{\underline{k}}}{\partial \underline{k}} f_{\underline{k}}^- \quad (21)$$

Taking a variational derivative with respect to $f_{\underline{k}}^-$ gives

$$\sum_{\underline{k}} \underline{k} \delta f_{\underline{k}}^- = \sum_{\underline{k}} m \frac{\partial E_{\underline{k}}}{\partial \underline{k}} \delta f_{\underline{k}}^- + \sum_{\underline{k}, \underline{k}', \sigma'} m \frac{\partial}{\partial \underline{k}} \left(\frac{\delta E_{\underline{k}}}{\delta f_{\underline{k}'}^-} \right) f_{\underline{k}}^- \delta f_{\underline{k}'}^- \quad (22)$$

Since $\delta f_{\underline{k}}^-$ is arbitrary, one has, by using Eq. (10),

$$\frac{\underline{k}}{m} = \frac{\partial E_{\underline{k}}}{\partial \underline{k}} - \sum_{\underline{k}', \sigma'} \Phi_{\underline{k} \underline{k}'} \frac{\partial}{\partial \underline{k}'} f_{\underline{k}'}^- \quad (23)$$

For \underline{k} close to the Fermi surface and at sufficiently low temperatures, Eq. (23) reduces to

$$\frac{1}{m} = \frac{1}{m^*} - \frac{2k_F}{(2\pi)^3} \int d\Omega \Phi(k_F, \theta) \cos \theta \quad (24)$$

The Friedel Sum Rule

I want now to turn to a relation of another kind which provides some useful information about the potential experienced by a quasi-particle in the neighborhood of a static impurity embedded in the system. This is the Friedel sum rule which states that if an impurity of valence Z is introduced into the system, the potential experienced by a quasi-particle at absolute zero obeys the relation

$$Z = \frac{2}{\pi} \sum_{\ell} (2\ell + 1) \delta_{\ell}(k_F) \quad (25)$$

where $\delta_{\ell}(k_F)$ is the phase shift for the scattering of a quasi-particle at the Fermi surface with angular momentum ℓ by the screened impurity potential. The sum rule (25) was proved by Friedel for independent

electrons, but the proof is also valid in the Landau theory because it is based essentially only on Eq. (1). Choose for r the quantum numbers of a free electron in a large spherical box. These are the angular momentum quantum numbers ℓ and m , the spin σ , and the radial wave number $k_{\ell, \lambda}$ which is restricted by the boundary condition to take on values such that

$$j_{\ell}(k_{\ell, \lambda} R) = 0 \quad (26)$$

where R is the radius of the box and j_{ℓ} is the spherical Bessel function of order ℓ that is regular at the origin. From the asymptotic form

$$j_{\ell}(kR) \sim \sin\left(kR - \frac{\ell\pi}{2}\right) \quad kR \gg \frac{\ell\pi}{2}$$

note that

$$k_{\ell, \lambda} \approx \frac{\pi}{R} \left(\lambda + \frac{\ell}{2} \right)$$

with λ an integer. Now if an impurity is introduced at the center of the box the quasi particle energy levels will shift. One can think of them as now being the eigenstates of a model Hamiltonian which is the sum of that for the free system and that for the screened impurity potential. For large R the allowed values of k are now

$$k \approx \frac{1}{R} \left(\pi\lambda + \frac{\ell\pi}{2} - \delta_{\ell}(k) \right) \quad (27)$$

where $\delta_{\ell}(k)$ is the phase shift for the scattering of a quasi particle with angular momentum ℓ and momentum k by the screened impurity potential. The additional number of quantum states of angular momentum ℓ that are introduced into the interval $k_1 < k < k_2$ by the perturbation is

$$\Delta N_{\ell} = \frac{1}{\pi} [\delta_{\ell}(k_2) - \delta_{\ell}(k_1)] \quad (28)$$

The total number of states introduced in this interval is thus

$$\begin{aligned} \Delta N &= \sum_{\ell, m, \sigma} \frac{\delta_{\ell}(k_2) - \delta_{\ell}(k_1)}{\pi} \\ &= \frac{2}{\pi} \sum_{\ell} (2\ell + 1) [\delta_{\ell}(k_2) - \delta_{\ell}(k_1)] \end{aligned} \quad (29)$$

The phase shifts introduced in (27) contain an arbitrary multiple of π . One can remove this arbitrariness by requiring that $\delta_{\ell} \rightarrow 0$ as the potential

is removed. Then it is well known that

$$\delta_{\ell}(0) = n_{\ell} \pi \quad (30)$$

where n_{ℓ} is the number of bound states of angular momentum ℓ . With this choice, the total number of additional quantum states (both bound states and scattering states) introduced by the perturbation below momentum k_F is

$$N = \frac{2}{\pi} \sum_{\ell} (2\ell + 1) \delta_{\ell}(k_F) \quad (31)$$

Now if the system is at constant chemical potential, the Fermi momentum is unchanged and a number of electrons given by (31) will flow into the system. To complete the derivation of (25) one must argue that if one introduces into a system of a given number of particles an impurity of valence Z and also Z electrons, the net effect on the system must have an atomic range. Thus the shift in the chemical potential will be of order $(\text{volume})^{-1}$ and negligible. The calculation (31) at fixed chemical potential is, therefore, applicable to the case where $N = Z$. Equation (25) follows.

Transport Properties

In the Landau theory transport properties are calculated using a Boltzmann equation for the distribution function. We shall consider only two simple transport properties, namely, the electrical and thermal resistances due to impurities. These properties are particularly easy to calculate because the scattering mechanism is external to the system and does not depend on interactions between quasi particles. As a result the calculation reduces formally to the customary derivation for independent electrons. All that is needed to make the derivation apply to the Landau theory is a different interpretation of the single-particle energy spectrum. We now sketch the calculation, largely for completeness.

The Boltzmann equation has the form

$$\left(\frac{\partial f^-}{\partial t} \right)_{\text{driving}} + \left(\frac{\partial f^-}{\partial t} \right)_{\text{collision}} = 0 \quad (32)$$

For scattering by a dilute distribution of impurities, the second term is

$$\left(\frac{\partial f_{\mathbf{k}}^-}{\partial t} \right)_{\text{collision}} = \sum_{\mathbf{k}'} \omega(\mathbf{k}, \mathbf{k}') [f_{\mathbf{k}}^-, f_{\mathbf{k}'}^+ - f_{\mathbf{k}}^+, f_{\mathbf{k}'}^-] \quad (33)$$

$$= \sum_{\underline{k}'} \omega(\underline{k}, \underline{k}') (f_{\underline{k}'}^- - f_{\underline{k}}^-)$$

Here $\omega(\underline{k}, \underline{k}')$ is the product of the concentration of impurities and the scattering rate due to the same screened impurity potential that occurred in the Friedel sum rule. For spherical symmetry and elastic scattering,

$$\omega(\underline{k}, \underline{k}') = \omega(k, \Theta) \delta(k - k') = \sum_{\ell} \omega_{\ell} P_{\ell}(\cos \Theta) \delta(k - k') \quad (34)$$

where $\cos \Theta = \hat{k} \cdot \hat{k}'$, and P_{ℓ} is the Legendre polynomial. Making for the distribution function the Ansatz:

$$f_{\underline{k}}^- = f_{\underline{k}}^-(T) + \delta f_{\underline{k}}^- \quad (35)$$

$$\delta f_{\underline{k}}^- = f_{\ell m}(k) Y_{\ell m}(\theta, \varphi)$$

One finds by using the spherical harmonic addition theorem that

$$\left(\frac{\partial f_{\underline{k}}^-}{\partial t} \right)_{\text{collision}} = -\frac{\delta f_{\underline{k}}^-}{\tau(k)} \quad (36)$$

$$\frac{1}{\tau(k)} = \sum_{\underline{k}'} \omega(\underline{k}, \underline{k}') (1 - P_{\ell}(\cos \Theta))$$

In an electrical resistance measurement the driving force is an electric field \underline{E} . The driving term in the Boltzmann equation is

$$\left(\frac{\partial f_{\underline{k}}^-}{\partial t} \right)_{\text{driving}} = -\frac{\partial f_{\underline{k}}^-(T)}{\partial \underline{k}} \cdot e \underline{E} = -\frac{\partial f_{\underline{k}}^-(T)}{\partial E_k} e \frac{\partial E_k}{\partial \underline{k}} \cdot \underline{E} \quad (37)$$

The electrical current is then

$$\underline{j} = e \sum_{\underline{k}, \sigma} \frac{\partial E_k}{\partial \underline{k}} \delta f_{\underline{k}}^- = -\frac{2e^2}{3} \sum_{\underline{k}} v_k^2 \frac{\partial f_{\underline{k}}^-(T)}{\partial E_k} \tau(k) \underline{E} \quad (38)$$

with $\underline{v}_k = \partial E_k / \partial \underline{k}$ and τ given by (37) with $\ell = 1$. For the conductivity at absolute zero one obtains the well-known result

$$\sigma = \frac{ne^2 \tau}{m^*} \quad (39)$$

where n the number density of electrons and τ is the relaxation time for a P_1 distortion of the Fermi surface. In terms of the differential scattering cross section $d\sigma/d\Omega$ for an electron at the Fermi surface, one has (n_i is the density of impurities)

$$\frac{1}{\tau} = n_i v_F \int d\Omega \frac{d\sigma}{d\Omega} (1 - \cos \theta) \quad (40)$$

Finally to calculate the heat conductivity one has to calculate the energy current that flows in a temperature gradient when no mass current is allowed. From the Boltzmann equation

$$f_k^- - f_k^-(T) = \tau(k) \left[- \frac{\partial f_k^-(T)}{\partial T} v_k \frac{\partial T}{\partial x} - e \frac{\partial f_k^-(T)}{\partial E_k} \frac{\partial E_k}{\partial x} E \right] \quad (41)$$

and one arrives at the set of equations

$$\begin{aligned} j &= -\frac{\mu}{e} L_0 \left(\frac{1}{\mu} \frac{\partial \mu}{\partial x} - \frac{1}{T} \frac{\partial T}{\partial x} \right) - L_1 \frac{1}{T} \frac{\partial T}{\partial x} \\ Q &= -\frac{\mu}{e} L_1 \left(\frac{1}{\mu} \frac{\partial \mu}{\partial x} - \frac{1}{T} \frac{\partial T}{\partial x} \right) - L_2 \frac{1}{T} \frac{\partial T}{\partial x} \end{aligned} \quad (42)$$

where j is the electrical current and Q the heat current given by

$$Q = 2 \sum_k E_k v_k f_k^- \quad (43)$$

Above

$$\begin{aligned} L_i &= -\frac{2e^2}{3e^4} \sum_k E_k^i v_k^2 \tau(k) \frac{\partial f_k^-(T)}{\partial E_k} \\ &= -\frac{1}{e^4} \int dE \frac{\partial f_k^-(T)}{\partial E} E^i \sigma(E) \quad i = 0, 1, 2 \end{aligned} \quad (44)$$

In (44), $\sigma(\mu)$ is the d-c conductivity at $T = 0$ [Eqs. (39) and (40)] that occurs here because an electric field in the absence of concentration and temperature gradients is given by $E = (1/e) \nabla \mu$. In a heat conduction experiment $j = 0$ and

$$Q = \frac{1}{T} \left(\frac{L_1^2}{L_0} - L_2 \right) \frac{\partial T}{\partial x} = -\kappa \frac{\partial T}{\partial x} \quad (45)$$

where κ is the thermal conductivity. To lowest order in temperature,

$$1_{\frac{1}{2}} = \frac{1}{\omega^2} \omega^2 = 0 \quad (46)$$

and κ is zero. However the first temperature correction in the Sommerfeld expansion of f , i.e.,

$$f(T) = f(E - \mu) + \frac{k_B^2 T^2}{6} \left(\frac{\partial^2}{\partial E^2} \right) f(E - \mu) \quad (47)$$

leads to a nonzero result and one easily finds

$$\kappa = \frac{\pi^2 k_B^2 T}{3e^2} \quad (48)$$

The proportionality between κ and σ is called the Wiedemann-Franz law. Experimentally this law is well obeyed by the residual or impurity dominated parts of these quantities.

This concludes my discussion of the Landau theory. We now turn to the microscopic theory to see to what extent the above results can be justified from a more fundamental starting point. In outline the program is as follows. In Sec. 2 we discuss the formal properties of Green's functions and show what assumptions are needed in order to get normal quasi-particles. In Sec. 3 we begin with a derivation of the Friedel sum rule, and then derive the equilibrium properties. Section 4 is concerned with the transport properties. Appendices A, B, and C supply some proofs that are missing in the main text. In Appendix D a proof of the effective mass sum rule (24) is given.

2. FORMAL PROPERTIES OF GREEN'S FUNCTIONS

We shall consider a system of fermions described by the Hamiltonian

$$H = \sum_i \frac{p_i^2}{2m} + \sum_{i < j} v(x_i - x_j) \quad (49)$$

assuming the interaction to be spin independent. In the language of second quantization

$$H = - \int d\tau \psi_{\alpha}^{\dagger}(\underline{x}) \frac{\nabla^2}{2m} \psi_{\alpha}(\underline{x}) + \frac{1}{2} \int d\tau d\tau' v(\underline{x} - \underline{x}') \times \psi_{\alpha'}^{\dagger}(\underline{x}) \psi_{\alpha'}^{\dagger}(\underline{x}') \psi_{\alpha'}(\underline{x}') \psi_{\alpha}(\underline{x}) \quad (50)$$

where a sum over spin states is implicit in the volume integrals. The field operators obey the usual fermion anticommutation relations

$$\begin{aligned} \{\psi_{\alpha}(\underline{x}), \psi_{\alpha'}^{\dagger}(\underline{x}')\} &= \delta_{\alpha\alpha'} \delta(\underline{x} - \underline{x}') \\ \{\psi_{\alpha}(\underline{x}), \psi_{\alpha'}(\underline{x}')\} &= \{\psi_{\alpha}^{\dagger}(\underline{x}), \psi_{\alpha'}^{\dagger}(\underline{x}')\} = 0 \end{aligned} \quad (51)$$

We shall usually omit the spin index α in the future. The number operator is given by

$$N = \int \sum_i \delta(\underline{x}_i - \underline{x}) d\tau = \int d\tau \psi^{\dagger}(\underline{x}) \psi(\underline{x}) \quad (52)$$

In calculating the properties of the system described by (1), we shall need to know how the system, in contact with a thermal reservoir, responds to disturbances of various kinds. To obtain the single-particle excitation spectrum of the system we require its response to the addition or removal of a single particle. To calculate the transport properties we shall need to know the response to the excitation of a single particle or the creation of a particle hole pair. The quantities that contain information of this sort are the thermodynamic Green's functions. These are defined to be expectation values of time-ordered field operators in the grand canonical ensemble. One uses the grand canonical ensemble simply for computational convenience, it being known from statistical mechanics that for large systems the various ensembles are equivalent. The single-particle Green's function is defined for times $t > t'$ according to

$$G(\underline{x}, t; \underline{x}', t') = - \frac{i \text{Tr} \left\{ e^{-\beta(H - \mu N)} \psi(\underline{x}, t) \psi^{\dagger}(\underline{x}', t') \right\}}{\text{Tr} \left\{ e^{-\beta(H - \mu N)} \right\}} \quad (53)$$

where the field operators $\psi(\underline{x}, t)$, $\psi^{\dagger}(\underline{x}', t')$ are in the Heisenberg representation

$$\psi(\underline{x}, t) = e^{iHt} \psi(\underline{x}) e^{-iHt}$$

For $t < t'$

$$G(\underline{x}, t; \underline{x}', t') = \frac{i \operatorname{Tr} \left\{ e^{-\beta(H-\mu N)} \psi^+(\underline{x}', t') \psi(\underline{x}, t) \right\}}{\operatorname{Tr} \left\{ e^{-\beta(H-\mu N)} \right\}} \quad (54)$$

It is convenient to write for the trace operations

$$\langle X \rangle = \frac{\operatorname{Tr} \left\{ e^{-\beta(H-\mu N)} X \right\}}{\operatorname{Tr} \left\{ e^{-\beta(H-\mu N)} \right\}} \quad (55)$$

The operator

$$\frac{e^{-\beta(H-\mu N)}}{\operatorname{Tr} \left\{ e^{-\beta(H-\mu N)} \right\}} \quad (56)$$

is just the density matrix for a given temperature and chemical potential.

In the notation (55) and using the shorthand $(1) = (\underline{x}, t)$, $(1') = (\underline{x}', t')$, we have

$$G(1, 1') = -i T \langle \psi(1) \psi^+(1') \rangle \quad (57)$$

Here T is Wick's time-ordering operator, which orders the operators so that the later time comes to the left, and gives a total sign equal to the sign of the permutation of the operators that gives the correct time ordering. The single-particle Green's function describes for positive time differences the propagation of an additional particle which is inserted in the medium. Specifically, for $t > t'$, $G(\underline{x}, t; \underline{x}', t')$ is the probability amplitude for a particle inserted at (\underline{x}', t') to be found at a later time t at the point (\underline{x}) . For $t < t'$, the Green's function in a similar way describes the propagation of a hole.

The two-particle Green's function is defined by

$$G(1, 2; 1', 2') = (-i)^2 T \langle \psi(1) \psi(2) \psi^+(2') \psi^+(1') \rangle \quad (58)$$

In general the n -particle Green's function is defined by

$$G(1, 2, \dots, n; 1', 2', \dots, n') = (-i)^n T \langle \psi(1) \psi(2) \dots \psi(n) \psi^+(n') \dots \psi^+(1') \rangle \quad (59)$$

Theory of the Single-Particle Green's Function

Considerable information about the single-particle Green's function can be obtained just from its definition (57). By studying the structure of this form we shall be able to learn the easiest ways of obtaining the information contained in it. As a start, let us consider a noninteracting system. For such a system the traces in (53) and (59) are easy to evaluate. We expand the field operators in terms of the complete set of orthonormal single-particle eigenstates of the Hamiltonian

$$\psi(\underline{x}, t) = \sum_r u_r e^{-i\epsilon_r t} a_r \quad (60)$$

The single-particle Green's function is then

$$G(\underline{x}, t; \underline{x}', t') = \begin{cases} -i \sum_{rs} u_r(\underline{x}) u_s^*(\underline{x}') e^{-i\epsilon_r t} e^{i\epsilon_s t'} \langle a_r a_s^+ \rangle, & t' < t \\ i \sum_{rs} u_r(\underline{x}) u_s^*(\underline{x}') e^{-i\epsilon_r t} e^{i\epsilon_s t'} \langle a_s^+ a_r \rangle, & t' > t \end{cases} \quad (61)$$

The expectation values in (61) are familiar from statistical mechanics. We have

$$\langle a_s^+ a_r \rangle = \delta_{rs} f_r^- = \left(e^{\beta(\epsilon_r - \mu)} + 1 \right)^{-1} \quad (62)$$

and

$$\langle a_r a_s^+ \rangle = \delta_{rs} (1 - f_r^-) = \delta_{rs} f_r^+ = \left(1 + e^{-\beta(\epsilon_r - \mu)} \right)^{-1} \quad (63)$$

By substituting (62) and (63) into (61), we get

$$G(\underline{x}, t; \underline{x}', t') = \begin{cases} -i \sum_r u_r(\underline{x}) u_r^*(\underline{x}') e^{-i\epsilon_r(t-t')} f_{r,t}^+, & t > t' \\ i \sum_r u_r(\underline{x}) u_r^*(\underline{x}') e^{-i\epsilon_r(t-t')} f_{r,t}^-, & t < t' \end{cases} \quad (64)$$

In (64) we see a single oscillatory time dependence associated with each single-particle level of the system. In general, the single-particle Green's function will give information about the existence and energies of single-particle-like excitations. For an interacting system, the Green's function will by no means be of the simple form (64) but it may be that there is a definite energy associated with, for example, a given momentum. If this is so, we shall see that the time dependence gives the energy of a

single-particle-like excitation with, say, momentum p . In other words, the single-particle excitation energies will be associated with the singularities in the frequency plane of the Fourier transform, with respect to time, of the single-particle Green's function.

Let us now consider the general case. We shall see that an examination of the formal properties of the Green's function will be greatly facilitated by a continuation to complex time. We write

$$G(\underline{x}, t; \underline{x}', t') = G_{>}(\underline{x}, t; \underline{x}', t') \theta(t - t') + G_{<}(\underline{x}, t; \underline{x}', t') \theta(t' - t) \quad (64)$$

and use the cyclic property of traces,

$$\text{Tr} \{ABC\} = \text{Tr} \{CAB\} = \text{Tr} \{BCA\} \quad (65)$$

to get

$$G_{>}(\underline{x}, t; \underline{x}', t') = -i \frac{\text{Tr} \left\{ e^{-\beta(H - \mu N)} e^{iH(t-t')} \psi(\underline{x}) e^{-iH(t-t')} \psi^+(\underline{x}') \right\}}{\text{Tr} \left\{ e^{-\beta(H - \mu N)} \right\}} \quad (66)$$

and

$$G_{<}(\underline{x}, t; \underline{x}', t') = i \frac{\text{Tr} \left\{ e^{-\beta(H - \mu N)} \psi^+(\underline{x}') e^{iH(t-t')} \psi(\underline{x}) e^{-iH(t-t')} \right\}}{\text{Tr} \left\{ e^{-\beta(H - \mu N)} \right\}} \quad (67)$$

Here we have used the fact that

$$[H, N] = 0 \quad (68)$$

and we see that

$$G(t, t') = G(t - t') \quad (69)$$

It is natural to assume that the traces in (66) and (67) converge absolutely whenever the coefficients of H in the exponentials have negative real parts. With this assumption we see that with reference to Fig. 1,

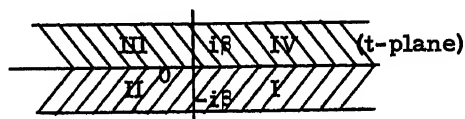


Fig. 1

$G_>$ may be analytically continued to the regions I and II in the time difference plane, and $G_<$ to the regions III and IV. We, therefore, extend the definition of the Green's function as follows:

$$G(\underline{x}, t; \underline{x}', t') = \begin{cases} G_>(\underline{x}, t; \underline{x}', t') & (t - t') \text{ in I} \\ G_<(\underline{x}, t; \underline{x}', t') & (t - t') \text{ in III} \end{cases} \quad (70)$$

In the extended region of the time difference, $G_>$ and $G_<$ are not independent. For, using the relation

$$\psi^+ g(N+1) = g(N) \psi^+ \quad (71)$$

and the cyclic relation (65), it is easily seen that for t in regions III or IV

$$G_>(t - i\beta) = -e^{\beta\mu} G_<(t) \quad (72)$$

and, similarly, for t in regions I and II

$$G_<(t + i\beta) = -e^{\beta\mu} G_>(t) \quad (73)$$

The periodicity conditions (72) and (73) can be used to construct a spectral representation for G . Making Fourier transformations we have

$$G_>(\omega) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} G_>(\omega) e^{-i\omega t} \quad t \text{ in I U II} \quad (74)$$

and

$$G_<(\omega) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} G_<(\omega) e^{-i\omega t} \quad t \text{ in III U IV} \quad (75)$$

we see from (72) and (73) that

$$G_>(\omega) = -e^{-\beta\mu} e^{\beta\omega} G_<(\omega) \quad (76)$$

The spectral weight function $A(\omega)$ is defined according to

$$G_>(\omega) = -i A(\omega) f^+(\omega) \quad (77)$$

It follows from (76) that

$$G_<(\omega) = i A(\omega) f^-(\omega) \quad (78)$$

Here

$$f^{\pm}(w) = (e^{\mp\beta(w-\mu)} + 1)^{-1} \quad (79)$$

Now, for real t

$$G(t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} G(\omega) e^{-i\omega t}$$

and by using the integral representation for the step function

$$\theta(t) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{e^{i\omega t} d\omega}{\omega - i\eta} \quad (80)$$

we get

$$G(\omega) = i \int_{-\infty}^{\infty} \frac{d\bar{\omega}}{2\pi} \left(\frac{G_{>}(\bar{\omega})}{\omega - \bar{\omega} + i\eta} - \frac{G_{<}(\bar{\omega})}{\omega - \bar{\omega} - i\eta} \right) \quad (81)$$

Substituting (77) and (78) into (81):

$$G(\omega) = \int_{-\infty}^{\infty} \frac{d\bar{\omega}}{2\pi} A(\bar{\omega}) \left(\frac{f^{+}(\bar{\omega})}{\omega - \bar{\omega} - i\eta} + \frac{f^{-}(\bar{\omega})}{\omega - \bar{\omega} - i\eta} \right) \quad (82)$$

Equation (82) was first written down by Landau.⁽⁵⁾ At zero temperature the Fermi factors reduce to step functions and we get the Migdal spectral representation.⁽⁶⁾

It is important to note that the single-particle Green's function is completely determined once $A(\omega)$ is known. This, therefore, is an important quantity and we shall now discuss its properties and interpret it physically. Consider a translationally invariant system. By Fourier transforming with respect to the coordinate difference, we have from the definitions (77) and (78)

$$\begin{aligned} A(\underline{p}, \omega) &= i [G_{>}(\underline{p}, \omega) - G_{<}(\underline{p}, \omega)] \\ &= \sum_{nm} \rho_n [\langle n | a_{\underline{p}} | m \rangle \langle m | a_{\underline{p}}^{\dagger} | n \rangle 2\pi \delta(E_n - E_m + \omega) \\ &\quad + \langle n | a_{\underline{p}}^{\dagger} | m \rangle \langle m | a_{\underline{p}} | n \rangle 2\pi \delta(E_m - E_n + \omega)] \end{aligned} \quad (83)$$

where $|n\rangle$ and $|m\rangle$ are eigenstates of the many particle Hamiltonian H and ρ_n is the (diagonal) density matrix in this representation. From (83), we

see that, at a given temperature, $A(\underline{p}, \omega)$ is the equilibrium density of states that can be reached by adding or removing a particle of momentum \underline{p} and energy ω in the interacting system.

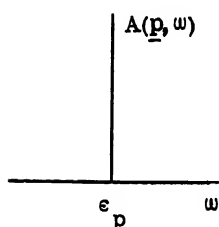


Fig. 2. Non-interacting system

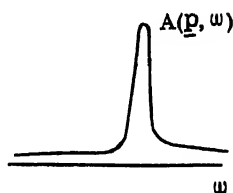


Fig. 3. Normal fermion system

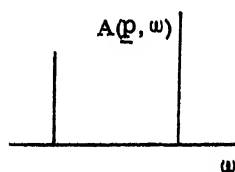


Fig. 4. BSC theory

For a noninteracting system, $A(\underline{p}, \omega)$ has a very simple form (see Fig. 2). It is easy to see from (83) or equivalently from the form of the Green's function (64) that

$$A(\underline{p}, \omega) = 2\pi \delta(\omega - \epsilon_{\underline{p}}) \quad (84)$$

For an interacting system, $A(\underline{p}, \omega)$ will not have any such simple form and, in general, will depend on the temperature. However, if $A(\underline{p}, \omega)$ for a given \underline{p} is strongly peaked for some value of ω , this indicates a single-particle-like excitation of the system with energy ω . We shall see that, for a normal Fermi system, $A(\underline{p}, \omega)$ has just such a behavior for \underline{p} near a certain value (Fermi momentum) (see Fig. 3). On the other hand, we may mention that for the BCS theory of superconductivity (which we shall not discuss at all here), $A(\underline{p}, \omega)$ has two peaks in ω for certain values of \underline{p} , corresponding to the fact that the quasi particles in superconductors are mixtures of particle and hole states (see Fig. 4).

There are two properties of $A(\underline{p}, \omega)$ that are quite general. These follow from the explicit form (83) and from the basic commutation relations of the fermion field operators. From (83) we see that

$$A(\underline{p}, \omega) \geq 0 \quad (85)$$

Furthermore, we have

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} A(\underline{p}, \omega) &= i [G_{>}(\underline{p}, t=0) - G_{<}(\underline{p}, t=0)] \\ &= \langle a_{\underline{p}}^{\dagger} a_{\underline{p}} + a_{\underline{p}}^{\dagger} a_{\underline{p}} \rangle = 1 \end{aligned} \quad (86)$$

Having seen that the single-particle Green's function is completely determined by $A(\underline{p}, \omega)$, we want to show how this latter quantity may most easily be extracted from a microscopic theory. Referring to Fig. 1, we see that the imaginary axis of the time difference variable from $-i\beta$ to $i\beta$ plays a particularly important role. Along this line the periodicity conditions (72) and (73) connect the Green's function (70) with itself. Because of this it is always easiest, as we shall see later, to compute the Green's function along this line. If we restrict all times to the region (hereafter called the "restricted region")

$$\operatorname{Re} t = 0 \quad -\beta < \operatorname{Im} t < 0 \quad (87)$$

the time differences are confined to the region mentioned. We define

$$\tau = -i\beta \quad (88)$$

and write for the restricted region $0 < t, t' < \tau$, the inequality sign referring to the negative imaginary part.[†] With these conventions the periodicity conditions (72) and (73) for t, t' in the restricted time region read

$$\begin{aligned} G(0, t') &= G_{<}(-t') = -e^{+\beta\mu} G_{>}(\tau - t') = -e^{\beta\mu} G(\tau, t') \\ G(t, 0) &= G_{>}(t) = -e^{-\beta\mu} G_{<}(t - \tau) = -e^{-\beta\mu} G(t, \tau) \end{aligned} \quad (89)$$

It is natural to satisfy these conditions by periodically extending the Green's function in strips of width 2β parallel to the real axis, and making a Fourier series analysis

$$G(t - t') = \frac{1}{\tau} \sum_{\ell} G(\zeta_{\ell}) e^{-i\zeta_{\ell}(t-t')} \quad \ell = 0, \pm 1, \pm 2, \dots \quad (90)$$

where

$$\zeta_{\ell} = \mu + \frac{(2\ell + 1)\pi}{\tau} \quad (91)$$

To verify that (90) does, in fact, satisfy the periodicity conditions note that

$$\begin{aligned} G(t \pm \tau) &= \frac{1}{\tau} \sum_{\ell} G(\zeta_{\ell}) e^{\mp i\mu\tau} e^{\mp i(2\ell+1)\pi} e^{-i\zeta_{\ell}t} \\ &= -e^{\mp\beta\mu} \frac{1}{\tau} \sum_{\ell} G(\zeta_{\ell}) e^{-i\zeta_{\ell}t} = -e^{\mp\beta\mu} G(t) \end{aligned} \quad (92)$$

[†]This notation may seem confusing at first. The reason it is used is, of course, that $G_{>}$ is continued to the lower half-plane, whereas $G_{<}$ is defined only in the upper half-plane. We also warn the reader that t, t' might refer both to times and time differences and which is meant will be clear from the context.

In a microscopic theory it will be the Fourier series coefficients $G(\zeta_\ell)$ that are most easily obtained. To extract the physical information we need, however, to determine $G(t)$ for t along the real axis or, equivalently, as we have seen, to find $A(\omega)$. Let us now see how this may be done. Inverting the Fourier series (90) we have

$$G(\underline{p}, \zeta_\ell) = \int_0^\tau e^{i\zeta_\ell t} G(\underline{p}, t) dt = \int_0^\tau e^{i\zeta_\ell t} G_{>}(\underline{p}, t) dt \quad (93)$$

By substituting Eq. (77) into (90) we get

$$G(\underline{p}, \zeta_\ell) = \frac{1}{i} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} A(\underline{p}, \omega) f^+(\omega) \int_0^\tau e^{+i\zeta_\ell t} e^{-i\omega t} dt \quad (94)$$

which gives

$$G(\underline{p}, \zeta_\ell) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{A(\underline{p}, \omega)}{\zeta_\ell - \omega} \quad (95)$$

How can $A(\underline{p}, \omega)$ be inferred from (95)? Often this can be done by inspection. For the result of a calculation is to determine $G(\underline{p}, \zeta_\ell)$ not as a set of numbers but as a function evaluated at the points ζ_ℓ . One can, therefore, continue to the ζ plane and write

$$G(\zeta) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{A(\omega)}{\zeta - \omega} \quad (96)$$

It is seen that $G(\zeta)$ has a branch cut along the real axis, and that $A(\omega)$ is simply the discontinuity across the cut. Explicitly,

$$\begin{aligned} G(\lambda + i\eta) - G(\lambda - i\eta) &= \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} A(\omega) \left(\frac{1}{\lambda - \omega + i\eta} - \frac{1}{\lambda - \omega - i\eta} \right) \\ &= -2\pi i \int \frac{d\omega}{2\pi} A(\omega) \delta(\omega - \lambda) = \frac{1}{i} A(\lambda) \end{aligned} \quad (97)$$

There is, however, one difficulty in the procedure. When continuing $G(\zeta_\ell)$ to the complex ζ plane we overlooked an ambiguity. A function defined on the points ζ_ℓ cannot be continued in a unique way. For example the two functions

$$G(\zeta), \quad G\left(-e^{-i\pi(\zeta-\mu)} \zeta\right) \quad (98)$$

are identical at the points ζ_ℓ , but are completely different elsewhere. It can be shown⁽⁷⁾ that there is a unique analytical continuation of $G(\zeta_\ell)$ such that $G(\zeta)$ goes to zero along any straight line in the complex plane approaching infinity and is analytic off the real axis. It is just for this unique

continuation that we can go from (96) to (97). It is sometimes difficult to tell by inspection which continuation of the form (98) is the correct one. This is particularly true of calculations in perturbation theory in which, as we shall see, the expression for $G(\zeta_l)$ contains sums over intermediate variables ζ'_l . It is, therefore, necessary to learn how to do such sums explicitly. We shall illustrate how this may be done for a simple case. Suppose we are given

$$G(\underline{p}, \zeta_l) = -\frac{1}{\zeta_l - \epsilon_{\underline{p}}} \quad (99)$$

and asked to find $G(t)$ for real t . This is, of course, a perfectly trivial example since (99) holds for a noninteracting system, and we know the answer. We can also determine the answer by inspection using the method discussed above since we see at once that the discontinuity across the real axis of (99) is just given by Eq. (84), and $G(t)$ is then determined by (74) and (75). However, if we did not know these formal properties we should have to sum the series

$$G(\underline{p}, t) = \frac{1}{\tau} \sum_l G(\underline{p}, \zeta_l) e^{-i\zeta_l t} \quad (100)$$

for t in the restricted region. The easiest way to do this is to transform the sum into a contour integral surrounding the poles of a function that has simple poles at the values ζ_l . The functions $f^{\pm}(\zeta)$ defined according to

$$f^{\pm}(\zeta) = \left(e^{\pm i\tau(\zeta_l - i0)} + 1 \right)^{-1} \quad (101)$$

have the correct poles for this purpose, with residues $(\pm i\tau)^{-1}$. We can then sum the series (100) by integrating along the contour Γ (see Fig. 5),

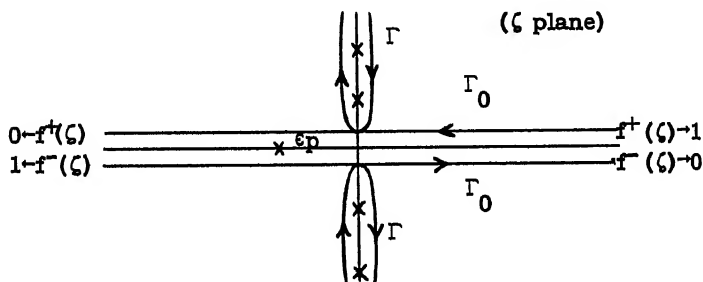


Fig. 5

$$G(\underline{p}, t) = \frac{1}{\tau} \frac{-1}{2\pi i} (\pm i\tau) \int_{\Gamma} d\zeta \frac{f^{\pm}(\zeta) e^{-i\zeta t}}{\zeta - \epsilon_{\underline{p}}} \quad (102)$$

If we choose the function (f^+ or f^-) for which the integrand is suitably convergent both to the left and to the right, we can deform the contour Γ into the contour Γ_0 without changing the value of the integral. Remembering that the inequality sign refers to the negative imaginary part and that we are considering t in the restricted time region, we see that the appropriate choice is f^+ for $t > 0$ and f^- for $t < 0$. The integral is then easily done since the contour Γ_0 contains only a single pole at $\zeta = \epsilon_{\underline{p}}$. We get

$$G(\underline{p}, t) = -i f^+(\epsilon_{\underline{p}}) e^{-i\epsilon_{\underline{p}} t} \quad t > 0 \quad (103)$$

$$G(\underline{p}, t) = i f^-(\epsilon_{\underline{p}}) e^{-i\epsilon_{\underline{p}} t} \quad t < 0$$

where t is in the restricted region. The extension to real t is now trivial, and one obtains the known result

$$G_{>}(\underline{p}, t) = -i f^+(\epsilon_{\underline{p}}) e^{-i\epsilon_{\underline{p}} t} \quad (t \text{ real}) \quad (104)$$

$$G_{<}(\underline{p}, t) = i f^-(\epsilon_{\underline{p}}) e^{-i\epsilon_{\underline{p}} t}$$

The above simple example illustrates a completely general method for doing sums over ζ_{ℓ} that will be of great use later in these lectures.

To summarize what has been said in this section: the quantity that completely determines the single-particle Green's function for real times is the spectral weight function $A(\omega)$. This function may be readily obtained from the Fourier series coefficient $G(\zeta_{\ell})$ for the Green's function on the imaginary time axis.

Before going on to show how the Fourier coefficients themselves may be calculated from a microscopic theory, we want to make some final remarks about the spectral weight function. We have already seen that a sharp peak in $A(\omega)$ indicates a single-particle-like excitation. What meaning can one give to a peak with a small but finite width? A good discussion of this point will be found in reference 6. We summarize that discussion here. Suppose that at $T = 0$, $A(\underline{p}, \omega)$ has the form

$$A(\underline{p}, \omega) = \frac{\Gamma_{\underline{p}}}{(\omega - E_{\underline{p}})^2 + \frac{\Gamma_{\underline{p}}^2}{4}} \quad (105)$$

where $\Gamma_{\underline{p}}$ is independent of ω and $\Gamma_{\underline{p}} < E_{\underline{p}}$. From (74) and (77) we have at $T = 0$

$$G_{>}(\underline{p}, t) = -i \int_{\mu}^{\infty} \frac{d\omega}{2\pi} A(\underline{p}, \omega) e^{-i\omega t} \quad (106)$$

If $E_p > \mu$ it is not difficult to show that for times $t > 1/\Gamma_p$, $G_{>}(t)$ has the form

$$G_{>}(t) = -i e^{-1\left(E_p - \frac{i\Gamma_p}{2}\right)t} + O\left(\frac{\Gamma_p}{E_p}\right) \quad (107)$$

The time $1/E_p$ thus has the interpretation of the "dressing time" or the time it takes for a bare particle introduced into the system to acquire its complement of virtual excitations and become a particle-like excitation of the many-particle system. The time $1/\Gamma_p$ is the decay time for the excitation once formed. The interpretation given above is strictly true only at $T = 0$. At finite temperatures the situation is complicated by the presence of thermal excitations. But we can still talk somewhat loosely of the width of the peaks in the spectral weight function as being a measure of the lifetime of the excitations.

Finally, it should be mentioned here that although the above analysis was concerned specifically with the single-particle Green's function, the same procedure can be applied to any correlation function that depends on one time difference. In particular, for the discussion of transport properties we shall have to consider the correlation function for two current operators. Much of the above analysis will be useful then with appropriate changes of interpretation.

Equations of Motion; Self-Energy Operator

We have now to consider the problem of calculating the Fourier coefficient $G(\zeta_{\lambda})$. This is of course a hard problem that can only be solved in some approximation. In this section we exhibit the formal structure of the problem by discussing the equations of motion for the Green's functions. These equations can and have been used to motivate approximation schemes. We shall not touch on this approach in the lectures. However, the formally exact discussion that follows will be useful later on in generating the perturbation expansions we shall use.

The field operators (in the Heisenberg representation) are governed by the equation of motion

$$i \frac{\partial}{\partial t} \psi(\underline{x}, t) = [\psi(\underline{x}, t), H] \quad (108)$$

Explicit computation of the commutator gives

$$i \frac{\partial}{\partial t} \psi(\underline{x}, t) = -\frac{\nabla^2}{2m} \psi(\underline{x}, t) + \int d^3x' v(\underline{x} - \underline{x}') \psi^+(\underline{x}', t') \psi(\underline{x}', t') \psi(\underline{x}, t) \quad (109)$$

This leads to the following equation of motion for the single-particle Green's function:

$$\begin{aligned} i \frac{\partial}{\partial t_1} G(\underline{x}_1, t_1; \underline{x}_2, t_2) &= \delta(t_1 - t_2) \langle \psi(\underline{x}_1, t_1) \psi^+(\underline{x}_2, t_1) \rangle \\ &+ \langle \psi^+(\underline{x}_1, t_1) \psi(\underline{x}_1, t_1) \rangle - \frac{\nabla_1^2}{2m} G(\underline{x}_1, t_1; \underline{x}_2, t_2) - i \int T \langle \psi^+(\underline{x}_1', t_1') \\ &\times \psi(\underline{x}_1', t_1') \psi(\underline{x}_1, t_1) \psi^+(\underline{x}_2, t_2) \rangle v(\underline{x}_1 - \underline{x}_1') d^3x_1' \end{aligned} \quad (110)$$

Here the index + (-) indicates a slightly later (earlier) time. We shall abbreviate (\underline{x}, t) , (\underline{x}', t') by (1), (1'), etc. and write

$$v(1, 1') = v(\underline{x} - \underline{x}') \delta(t - t') \quad (111)$$

Furthermore it is convenient to introduce a summation convention, namely that symbols with a bar over them should be integrated over all space and time. Using these conventions, Eq. (110) becomes

$$\left(i \frac{\partial}{\partial t_1} + \frac{\nabla_1^2}{2m} \right) G(1, 1') = \delta(1, 1') - i v(1, \bar{1}) G(1, \bar{1}; 1' \bar{1}^+) \quad (112)$$

For a noninteracting system, (112) reduces to

$$\left(i \frac{\partial}{\partial t_1} + \frac{\nabla_1^2}{2m} \right) G_0(1, 1') = \delta(1, 1') \quad (113)$$

Thus the differential operator in (112) may be interpreted as the inverse of the single-particle Green's function for a noninteracting system. As a result, (112) may be written

$$G_0^{-1}(1, \bar{1}) G(\bar{1}, 1') = \delta(1, 1') - i v(1, \bar{1}) G(1, \bar{1}; 1' \bar{1}^+) \quad (114)$$

Equation (114) is the first in an infinite hierarchy of equations that successively couple higher-order Green's functions. For the two-particle functions, we obtain

$$G_0^{-1}(1, \bar{1}) G(\bar{1}, 2; 1', 2') = \delta(1, 1') G(2, 2') - \delta(1, 2') G(2, 1') \\ - i v(1, \bar{1}) G(1, 2, \bar{1}; 1', 2', \bar{1}^+) \quad (115)$$

In deriving the integrodifferential equations above, the only property of the statistical average needed is the normalization property

$$\langle 1 \rangle = 1 \quad (116)$$

In order to obtain the thermodynamic Green's functions we have to supplement the differential equations by boundary conditions. The appropriate boundary condition is the periodicity along the imaginary time axis. If we require that all times lie in the restricted region discussed in the last section, it is easy to see that

$$G(t_1, \dots, t_{i-1}, 0, t_{i+1}, \dots, t_n; t_1', \dots, t_n') \\ = -G(t_1, \dots, t_{i-1}, \tau, t_{i+1}, \dots, t_n; t_1', \dots, t_n') e^{+\beta\mu} \\ \times G(t_1, \dots, t_n; t_1', \dots, t_{i-1}', 0, t_{i+1}', \dots, t_n') \\ = -G(t_1, \dots, t_n; t_1', \dots, t_{i-1}', \tau, t_{i+1}', \dots, t_n') e^{-\beta\mu} \quad (117)$$

The boundary conditions (117) may be incorporated into the differential equations by requiring all times to be in the restricted region and making a Fourier expansion for all the Green's functions according to

$$G(t_1, \dots, t_n; t_1', \dots, t_n') = \frac{1}{\tau^{2n}} \sum_{\{\ell, \ell'\}} e^{i(\zeta_{\ell_1} t_1 + \dots + \zeta_{\ell_n} t_n)} e^{-i(\zeta_{\ell'_1} t_1' + \dots + \zeta_{\ell'_n} t_n')} \\ \times G(\zeta_{\ell_1}, \dots, \zeta_{\ell_n}; \zeta_{\ell'_1}, \dots, \zeta_{\ell'_n}) \quad (118)$$

where, because of the invariance under a simultaneous translation of all coordinates, the right-hand side of (118) is zero unless

$$\zeta_{\ell_1} + \dots + \zeta_{\ell_n} = \zeta_{\ell'_1} + \dots + \zeta_{\ell'_n} \quad (119)$$

The coupled hierarchy of differential equations becomes a hierarchy of algebraic equations for the Fourier coefficients. In principle one can determine all the Fourier coefficients from these equations. In practice, the above method is only useful when the series is truncated at a certain point. Recent experience shows that some care has to be taken when doing this and that inconsistencies can result from a neglect of important higher-order terms.

While we are discussing formally exact properties of the Green's function, we should introduce the self-energy operator defined according to

$$\Sigma(1, 1') = G_0^{-1}(1, 1') - G^{-1}(1, 1') \quad (120)$$

By multiplying (120) on the right by G and comparing with the equation of motion (114), we find

$$\Sigma(1, \bar{1}) G(\bar{1}, 1') = i v(1, \bar{1}) G(1, \bar{1}; 1' \bar{1}') \quad (121)$$

From (120) we find that the Fourier series coefficients $G(\underline{p}, \zeta_l)$ may be expressed in the form

$$G(\underline{p}, \zeta_l) = \frac{1}{\zeta_l - \epsilon_{\underline{p}} - \Sigma(\underline{p}, \zeta_l)} \quad (122)$$

where we have used the known form (99) for the Fourier series coefficient of the noninteracting system. Equation (122) can be continued to the complex ζ plane according to

$$G(\underline{p}, \zeta) = \frac{1}{\zeta - \epsilon_{\underline{p}} - \Sigma(\underline{p}, \zeta)} \quad (123)$$

It follows from the form (123) that $\Sigma(\underline{p}, \zeta)$ has the same analytic properties as $G(\underline{p}, \zeta)$.⁽⁸⁾ Specifically $\Sigma(\underline{p}, \zeta)$ is analytic everywhere except on the real axis where it has a branch cut. Thus we can write†

$$\Sigma(\underline{p}, \zeta) = a_{\underline{p}} + \int_{-\infty}^{\infty} \frac{\Gamma_{\underline{p}}(w) dw}{\zeta - w} \frac{1}{2\pi} \quad (124)$$

Here $\Gamma_{\underline{p}}(w)$ is the discontinuity of Σ across the cut. Defining $\Delta_{\underline{p}}(w)$ according to

$$\Sigma(\underline{p}, w + i\delta) = \Delta_{\underline{p}}(w) + i \frac{\Gamma_{\underline{p}}(w)}{2} \quad (125)$$

†The constant $a_{\underline{p}}$ occurs because $\Sigma(\underline{p}, \zeta)$ unlike $G(\underline{p}, \zeta)$ may approach a constant value as $\zeta \rightarrow \infty$ in the upper or lower half-plane.

it follows that

$$\Sigma(\underline{p}, \omega - i\delta) = \Delta_{\underline{p}}(\omega) + i \frac{\Gamma_{\underline{p}}(\omega)}{2} \quad (126)$$

From the general definition of $A(\underline{p}, \omega)$ as the discontinuity across the cut of $G(\underline{p}, \omega)$ we get the relation

$$A(\underline{p}, \omega) = \frac{\Gamma_{\underline{p}}(\omega)}{(\omega - \epsilon_{\underline{p}} - \Delta_{\underline{p}}(\omega))^2 + \Gamma_{\underline{p}}^2(\omega)/4} \quad (127)$$

It follows, using (85) that

$$\Gamma_{\underline{p}}(\omega) > 0$$

Perturbation Theory

In our work we shall evaluate the Green's functions by infinite-order perturbation expressions. We shall see that in this way we are led more or less naturally to the theory of the normal Fermi system.

The essential theorem for perturbation theory is a generalization of Wick's theorem. In our context this states that for a noninteracting system the n -particle Green's function is related to the single-particle Green's function according to the relation,

$$G_0(1, \dots, n; 1' \dots n') = \begin{vmatrix} G_0(1, 1') & \dots & G_0(1, n') \\ \dots & \dots & \dots \\ G_0(n, 1') & \dots & G_0(n, n') \end{vmatrix} \quad (128)$$

where the right-hand side means the determinant of the square array. Equation (128) is easily proved by induction from the equations of motion. For the two-particle Green's function we have

$$G_0^{-1}(1, \bar{1})G_0(\bar{1}, 2; 1' 2') = \delta(1, 1')G_0(2, 2') - \delta(1, 2')G_0(2, 1') \quad (129)$$

Multiplying (129) with G_0 gives

$$G_0(1,2;1'2') = \begin{vmatrix} G_0(1,1') & G_0(1,2') \\ G_0(2,1') & G_0(2,2') \end{vmatrix} \quad (130)$$

By using (130) in the equation of motion for the three-particle Green's function, one proves the theorem (128) for $n=3$. In a similar way it is easy to see that if the theorem is true for $n=m$ it is also true for $n=m+1$. Note that in this proof we have made no use of the detailed form of the expectation value. The first proof of Wick's theorem for thermodynamic expectation values⁽⁷⁾ did make use of the detailed form. It was, therefore, considerably longer than the proof given above.

It is easiest to make the perturbation expansion in the restricted region of time. Further it is convenient to use an interaction representation:

$$\psi_I(\underline{x}, t) = e^{iH_0 t} \psi(\underline{x}) e^{-iH_0 t} \quad (131)$$

where

$$H = H_0 + V \quad (132)$$

and H_0 is the sum of one-particle Hamiltonians. In this way all dependence on the interaction can be expressed in terms of the Dyson U matrix:

$$U(t, t') = e^{iH_0 t} e^{-iH(t-t')} e^{-iH_0 t'} \quad (133)$$

which satisfies the group properties

$$U(t, t'') = U(t, t') U(t', t'') \quad (134)$$

$$U(t, t) = 1 \quad (135)$$

The U matrix obeys the equation of motion

$$i \frac{\partial}{\partial t} U(t, t') = V_I(t) U(t, t') \quad (136)$$

where

$$V_I(t) = e^{iH_0 t} V e^{-iH_0 t} \quad (137)$$

Integrating (136) we find the usual formal expression

$$U(\tau, 0) = T \exp \left(-i \int_0^\tau V_I(t) dt \right) \quad (138)$$

The only difference between (138) and the customary U matrix is that imaginary times are used here.

By a judicious insertion of factors 1 we get

$$i Z_G G(1, 1') = T \operatorname{Tr} \left\{ e^{\beta \mu N} e^{-\beta H_0} U(\tau, t) \right. \\ \left. \times \psi_I(\underline{x}, t) U(t, t') \psi_I^\dagger(\underline{x}', t') U(t', 0) \right\} \quad (139)$$

where Z_G is the grand partition function

$$Z_G = \operatorname{Tr} \left\{ e^{-\beta(H - \mu N)} \right\} \quad (140)$$

We indicate averaging in the unperturbed ensemble, but with the same chemical potential as the perturbed system, as

$$\langle X \rangle_0 = \frac{\operatorname{Tr} \left\{ e^{\beta \mu N} e^{-\beta H_0} X \right\}}{\operatorname{Tr} \left\{ e^{\beta(\mu N - H_0)} \right\}} \quad (141)$$

In this notation

$$Z_G = \langle U(\tau, 0) \rangle_0 Z_{G_0} \quad (142)$$

and Eq. (87) becomes (the time ordering also applying to the operators in the U matrix)

$$G(1, 1') = -i \frac{T \langle U(\tau, 0) \psi_I(1) \psi_I^\dagger(1') \rangle_0}{\langle U(\tau, 0) \rangle_0} \quad (143)$$

Similarly the two-particle Green's function can be written

$$G(1, 2; 1', 2') = (-i)^2 \frac{T \langle U(\tau, 0) \psi(1) \psi(2) \psi^\dagger(2') \psi^\dagger(1') \rangle_0}{\langle U(\tau, 0) \rangle_0} \quad (144)$$

where the subscript I for the interaction representation has been dropped.

For $U(\tau, 0)$ we have the expansion

$$U(\tau, 0) = 1 - i \int_0^\tau V(t) dt + \frac{(-i)^2}{2} T \int_0^\tau dt \int_0^\tau dt' V(t) V(t') + \dots \quad (145)$$

Further, using the notation we have introduced

$$-i \int_0^T V(t) dt = -\frac{i}{2} v(\bar{1}, \bar{1}') T [\psi^+(\bar{1}^+) \psi^+(\bar{1}', +) \psi(\bar{1}') \psi(\bar{1})] \quad (146)$$

From Eqs. (143), (145), and (146) we obtain the following perturbation series for the single-particle Green's function:

$$\begin{aligned} \frac{Z_G}{Z_{G_0}} G(1, 1') &= G_0(1, 1') - \frac{1}{2} v(\bar{1}, \bar{1}') \\ &\times T \langle \psi(1) \psi(\bar{1}) \psi(\bar{1}) \psi(\bar{1}^+) \psi^+(\bar{1}^+) \psi^+(1) \rangle_0 \\ &+ \frac{i}{2! 2^2} v(\bar{1}, \bar{1}') v(\bar{2}, \bar{2}') T \langle \psi(1) \psi(\bar{1}) \psi(\bar{1}') \psi(\bar{2}) \psi(\bar{2}') \psi^+(\bar{2}^+) \psi^+(\bar{2}^+) \\ &\cdot \psi^+(\bar{1}^+) \psi^+(\bar{1}^+) \psi^+(1) \rangle_0 + \dots \end{aligned} \quad (147)$$

This may be rewritten in the form

$$\begin{aligned} \frac{Z_G}{Z_{G_0}} G(1, 1') &= G_0(1, 1') + \frac{i}{2} v(\bar{1}, \bar{1}') G_0(1, \bar{1}, \bar{1}'; 1', \bar{1}^+, \bar{1}^+) \\ &+ \frac{i^2}{2! 2^2} v(\bar{1}, \bar{1}') v(\bar{2}, \bar{2}') G_0(1, \bar{1}, \bar{1}', \bar{2}, \bar{2}'; 1', \bar{1}^+, \bar{1}^+, \bar{2}^+, \bar{2}^+) \\ &+ \dots \end{aligned} \quad (148)$$

The higher-order unperturbed Green's functions may now be expanded in terms of the determinant (125). We see that this notation very soon gets out of hand, so what one does is to represent the terms of the expansion by graphs. We represent $G_0(1, 1')$ by a straight line with an arrow pointing from $1'$ to 1 , and $v(\bar{1}, \bar{1}')$ by a horizontal wavy line. The first-order term in (148) contains six terms which are obtained by expanding the 3×3 determinant. The graphs which represent these terms are illustrated in Fig. 6. Notice that there are two types of graphs. Figures 6(a) and (b) contain parts that are not linked to the line entering and leaving the graph. The other graphs are linked.[†]

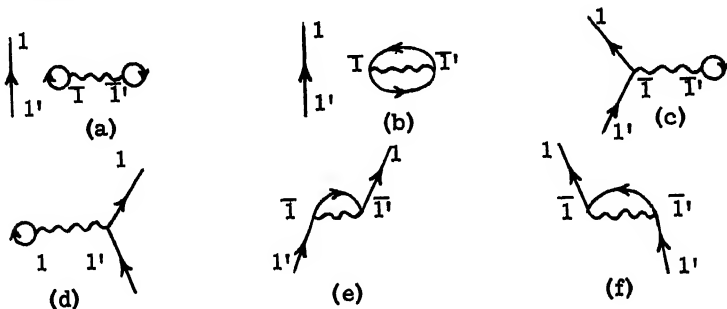


Fig. 6

From the nature of the derivation of the expansion (148) it is clear that to obtain the contribution of a particular graph with n interaction lines one has to multiply together the following factors:

- (1) $i^n/n!2^n$,
- (2) $v(\bar{i}, \bar{i}')$ for each interaction line,
- (3) $G(\bar{i}, \bar{j})$ for each particle line running from \bar{j} to \bar{i} .

Finally one has to associate a sign with each graph, which is the sign attached to the term that it represents in the original determinant. The rule for this sign as it is easy to see is

[†]Graphs (a), (c), and (d) contain closed loops that start and end at the same point. These loops describe the interaction with the static electronic charge distribution of the unperturbed system. We are actually interested in a neutral system in which the electronic charge is balanced by a background of positive charge. The only effect of such a background is to remove graphs such as those we are referring to here.

$$(4) \quad (-1)^c$$

where c is number of closed fermion loops of which, for example, Fig. 1(a) contains 2, Fig. 1(b), 1, and Fig. 1(e), none.

We have thus seen how to make an expansion of $(Z_G/Z_{G_0}) G(1, 1')$. It is easy to see that the factor Z_G/Z_{G_0} comes from the unlinked parts. Every linked graph has associated with it in higher order all possible unlinked graphs. Now if m is the order of the linked part, there are $2^m m!$ linked parts making an identical contribution which can be obtained by interchanging the end points of the interaction lines and permuting them among themselves. If we agree not to distinguish graphs that differ only in the labeling of the linked part we must multiply the contribution of any graph by $2^m m!$. If the combinatorial factor that remains is associated with the unlinked part one finds just the contribution one would get from an expansion of $\langle U(\tau, 0) \rangle_0$. Then using (142) one sees that $G(1, 1')$ is given by the sum of the contributions of the linked graphs above.

In making an expansion of the single-particle Green's function, we have, of course, also made an expansion of the self-energy operator Σ . From (120) it follows that

$$G = G_0 + G_0 \Sigma G_0 + G_0 \Sigma G_0 \Sigma G_0 + \dots \quad (149)$$

Comparing (149) with the perturbation expansion for G , one finds that Σ may be expressed as the sum of the contributions of all proper self-energy graphs. A self-energy graph is defined as a graph for G with the two terminal G_0 's removed. A proper self-energy graph cannot be broken into two parts by cutting one particle line. Figure 7 indicates a proper self-energy graph of fourth order, whereas Figure 8 indicates an improper graph of the same order.

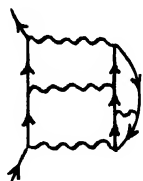


Fig. 7



Fig. 8

All the graphs for Σ share a property that is the crucial property of the normal Fermi system. This is that the spectral weight function of Σ , namely Γ [see Eq. (124)], at sufficiently low temperatures and for frequencies near the chemical potential behaves in leading order according to

$$\Gamma_p(\omega) = 2 C_p [(\omega - \mu)^2 + \pi^2 k_B^2 T^2] + \dots \quad (150)$$

The proof that this is generally so is most easily given using some powerful techniques we shall develop later. Therefore we defer the proof to Appendix A. We see from (150) using (125) that $\Sigma(\underline{p}, \mu)$ is real at $T = 0$

$$\Sigma(\underline{p}, \mu) = \Delta(\underline{p}) \quad (151)$$

It is often stated that (151) leads directly to normal Fermi properties. This does not appear to be so. Consider the equation.

$$\mu = \epsilon_{\underline{p}} + \Delta_{\underline{p}}(\mu) \quad (152)$$

For the normal Fermi system it is necessary that the right-hand side considered as a function of \underline{p} intersect the straight line μ once and only once. Now for the BCS theory, $\Sigma(\underline{p}, \mu)$ has a pole at $\epsilon_{\underline{p}} = \mu$. In particular it has the form

$$\Sigma(\underline{p}, \mu) = \frac{\Delta^2}{\epsilon_{\underline{p}} - \mu} \quad (153)$$

This is illustrated in Fig. 9. Now, although there is no reason to expect such singular behavior in

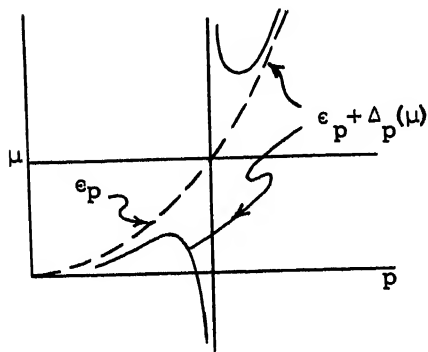


Fig. 9

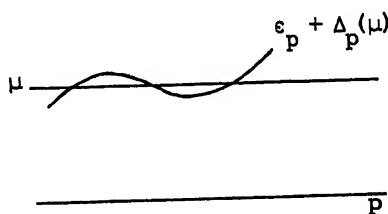


Fig. 10

perturbation theory, there also seems to be no reason for excluding the possibility indicated in Fig. 10. A system for which $\epsilon_{\underline{p}} + \Delta_{\underline{p}}(\mu)$ had this behavior would not be a normal Fermi system. It therefore seems necessary to assume that there is only one intersection. We call the momentum at this intersection the Fermi momentum p_F . One can then use a continuity argument to assert that for p near p_F the equation

$$\omega = \epsilon_{\underline{p}} + \Delta_{\underline{p}}(\omega) \quad (154)$$

will have a solution for p near p_F . The locus of the intersection (154) gives a curve which we shall call the quasi-particle dispersion curve E_p , for reasons that will become clear (see Fig. 11).

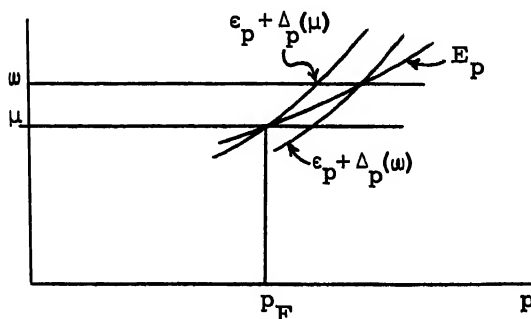


Fig. 11

We shall see below that a behavior of the self-energy operator of the type described leads to a system of normal quasi-particles. Thus we are led to such a system more or less naturally by the use of perturbation theory. In no sense can we be said to have answered the question of what the necessary and sufficient conditions on the two-body interaction $v(\underline{x}, \underline{x}')$ are in order that the results obtained from perturbation theory correspond to the true behavior. This is a difficult and, to my knowledge, an unsolved question. The only attitude we can take at this point is to assume that the above-mentioned conditions (whatever they are) are satisfied for the potentials we are considering. Some remarks about the philosophy behind this attitude will be made at the end of the lectures.

Let us now see what conclusions can be drawn about the single-particle Green's function from the properties of the self-energy operator assumed above. Quite generally we have

$$A(\underline{p}, \omega) = \frac{\Gamma_{\underline{p}}(\omega)}{[\omega - \epsilon_{\underline{p}} - \Delta_{\underline{p}}(\omega)]^2 + \frac{\Gamma_{\underline{p}}^2(\omega)}{4}} \quad (127)$$

Now, under the assumptions above

$$\omega - \epsilon_{\underline{p}} + \Delta_{\underline{p}}(\omega) \quad (155)$$

has a unique zero $\omega = E_{\underline{p}} \sim \mu$ for $p \sim p_F$. Let us therefore expand

$$\Delta_{\underline{p}}(\omega) = \Delta_{\underline{p}}(E_{\underline{p}}) + (\omega - E_{\underline{p}}) \left. \Delta'_{\underline{p}}(\omega) \right|_{\omega=E_{\underline{p}}} \quad (156)$$

and define

$$Z_{\underline{p}}^{-1} = 1 - \Delta'_{\underline{p}}(E_{\underline{p}}) \quad (157)$$

Measuring energies relative to μ , i.e., making the substitution

$$\omega - \mu \rightarrow \omega \quad E_{\underline{p}} - \mu \rightarrow E_{\underline{p}} \quad (158)$$

we have at $T = 0$ using (150)

$$A_{\underline{p}}(\omega) = \frac{2C_{\underline{p}}\omega^2}{Z_{\underline{p}}^{-2}(\omega - E_{\underline{p}})^2 + C_{\underline{p}}^2\omega^4} + A_{\underline{p}}^{(0)}(\omega) = A_{\underline{p}}^{(1)}(\omega) + A_{\underline{p}}^{(0)}(\omega) \quad (159)$$

where $A_{\underline{p}}^{(0)}(\omega)$ is a perfectly smooth function for $\omega \sim E_{\underline{p}}$ and $p \sim p_F$.

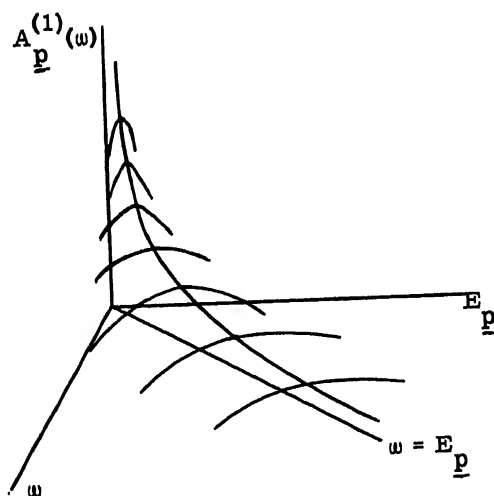


Fig. 12

$A_{\underline{p}}^{(1)}(\omega)$ as a function of ω for given \underline{p} is peaked near $\omega = E_{\underline{p}}$. The height of the peak is $\sim 2/C_{\underline{p}}\omega^2$ and the width $\sim 2Z_{\underline{p}}C_{\underline{p}}\omega^2$ (see Fig. 12). We see that as \underline{p} approaches the Fermi momentum, $A_{\underline{p}}^{(1)}(\omega)$ approaches more and more a δ function. Thus, at $T = 0$ we have

$$A_{\underline{p}}(\omega) \xrightarrow{\underline{p} \rightarrow \underline{p}_f} 2\pi Z_{\underline{p}} \delta(\omega - E_{\underline{p}}) + \tilde{A}_{\underline{p}}(\omega) \quad (160)$$

Note that from the sum rule (86)

$$0 < Z_{\underline{p}} \leq 1 \quad (161)$$

In the next section we shall see how the properties discussed here lead to the equilibrium results of the Landau theory.

3. EQUILIBRIUM PROPERTIES

In this section we shall use the single-particle Green's function derived above and the structure of perturbation theory to obtain the equilibrium properties of the Landau theory. The first property we shall discuss follows directly from the form of the spectral weight function (160), namely that the occupation probability in momentum space has a sharp discontinuity. In other words we find a many-particle analogue to the Fermi surface.

Let us count the number of particles with momentum \underline{p} :

$$\begin{aligned} n_{\underline{p}} &= \langle a_{\underline{p}}^+ a_{\underline{p}} \rangle = -i G(\underline{p}, t = 0^-) \\ &= -i \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} G_{<}(\underline{p}, \omega) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} A(\underline{p}, \omega) f^-(\omega) \end{aligned} \quad (162)$$

Thus as $T \rightarrow 0$

$$n_{\underline{p}} = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} A(\underline{p}, \omega) \quad (163)$$

Now using the form (160) we see that there is a discontinuity in momentum space given by

$$n_{p_F-0} - n_{p_F+0} = Z_{p_F} \quad (164)$$

So far the Fermi momentum has only been defined by the equation

$$\mu = \epsilon_{p_F} + \Delta_{p_F}(\mu) \quad (165)$$

The Fermi momentum p_{F0} for a noninteracting system is determined by the density of particles. We want now to show how p_F defined as above is related to p_{F0} . To do this we calculate the total number of particles in the system. One has

$$\begin{aligned} N &= -\frac{i}{\tau} \sum_{\underline{p}, \ell} G(\underline{p}, \zeta) e^{-i\zeta_\ell (i0^+)} \\ &= -\frac{i}{\tau} \sum_{\underline{p}, \ell} G(\underline{p}, \zeta_\ell) e^{\zeta_\ell 0^+} \end{aligned} \quad (166)$$

It is convenient at this point to use the identity

$$\begin{aligned} \frac{\partial}{\partial \zeta_\ell} \ln [\zeta_\ell - \epsilon_p - \Sigma(\underline{p}, \zeta_\ell)] \\ = G(\underline{p}, \zeta_\ell) \left(1 - \frac{\partial}{\partial \zeta_\ell} \Sigma(\underline{p}, \zeta_\ell) \right) \end{aligned} \quad (167)$$

By using this relation in (166) one obtains

$$\begin{aligned} N &= -\frac{i}{\tau} \sum_{\underline{p}, \ell} e^{\zeta_\ell 0^+} \frac{\partial}{\partial \zeta_\ell} \ln [\zeta_\ell - \epsilon_p - \Sigma(\underline{p}, \zeta_\ell)] \\ &\quad + \frac{i}{\tau} \sum_{\underline{p}, \ell} e^{\zeta_\ell 0^+} G(\underline{p}, \zeta_\ell) \cdot \frac{\partial}{\partial \zeta_\ell} \Sigma(\underline{p}, \zeta_\ell) \end{aligned} \quad (168)$$

This transformation is particularly useful because it can be shown that at $T = 0$ the second term goes to zero. This was shown by Luttinger⁽³⁾ using the topological structure of the graphs that occur in the perturbation expansion. We reproduce his proof in Appendix B. The vanishing of the second term in (168) is important for the argument that follows, and there must be a physical reason for it. We regret that we have not been able to discover this reason.

The sum over ζ_p in the first term of (168) may be done using the contour integral techniques discussed in the last section. We have (see Fig. 5), at $T = 0$,

$$N = \lim_{T \rightarrow 0} \frac{1}{2\pi i} \sum_{\underline{p}} \int_0^{\infty} e^{\zeta_0^+} f^-(\zeta) \times \frac{\partial}{\partial \zeta} \ln \left[\zeta - \epsilon_p - \Sigma(\underline{p}, \zeta) \right] d\zeta \quad (169)$$

If we choose the branch cut off the logarithm to run from $-\infty$ to 0, we have

$$N = \frac{1}{2\pi i} \sum_{\underline{p}} \left(\int_{-\infty - i\eta}^{\mu - i\eta} d\zeta - \int_{-\infty + i\eta}^{\mu + i\eta} d\zeta \right) \times \left(\frac{\partial}{\partial \zeta} \ln \left[(\zeta - \epsilon_p - \Sigma(\underline{p}, \zeta)) \right] \right) \quad (170)$$

which leads to

$$N = \sum_{\underline{p}} 1, \epsilon_p + \Sigma(\underline{p}, \mu) < \mu \quad (171)$$

In the noninteracting system the Fermi momentum and the number of particles are related (omitting the spin) by

$$\frac{N}{V} = \frac{4\pi}{3} p_{F0}^3 \frac{1}{(2\pi)^3} \quad (172)$$

For the interacting system we find the number of particles for a given μ by enumerating the eigenstates of the model single-particle Hamiltonian

$$h + \Delta(\mu) \quad (173)$$

with energy eigenvalues $\epsilon_p + \Delta_p(\mu) < \mu$. Under the assumptions we made use of earlier,

$$\epsilon_p + \Delta_p(\mu) < \mu \quad \text{for } p < p_F \quad (174)$$

Since the number of particles is not changed by the interaction this means that $p_F = p_{F0}$, i. e., the Fermi momentum is left invariant by the interaction.

In the above we have restricted ourselves to the translationally invariant case. This is largely for notational convenience. For example, if one had an additional static periodic potential, the appropriate single-particle quantum numbers would be a band index n and the quasi momentum \underline{k} . Most of the above analysis would still apply. In particular, one would still have a model Hamiltonian of the form (173), but this would in general not be diagonal in the original band indices. However, one can at least in principle diagonalize the model Hamiltonian and obtain the eigenvalues $\epsilon_{\underline{n}}(\underline{k})$ where \underline{n} are the new band indices. These bands would now fall into three classes, such that:

- I $\epsilon_{\underline{n}}(\underline{k}) = \mu$ is never satisfied and $\epsilon_{\underline{n}}(\underline{k}) < \mu$
- II $\epsilon_{\underline{n}}(\underline{k}) = \mu$ is never satisfied and $\epsilon_{\underline{n}}(\underline{k}) > \mu$
- III $\epsilon_{\underline{n}}(\underline{k}) = \mu$ is satisfied for some \underline{k}

The bands of class I correspond to filled bands, and those of class II to empty bands. The bands of class III contain the Fermi surface. In this general case of nonspherical Fermi surfaces one finds that the invariant quantity is the volume contained by the Fermi surface.

The Friedel Sum Rule

At this point it is very easy to give a many-particle derivation of the Friedel sum rule.⁽¹⁰⁾ We add to the system we have been considering a single impurity and wish to calculate the change in the number of electrons at fixed chemical potential. In analogy with (168) we have

$$\Delta N = -\frac{i}{\pi} \text{Tr} \sum_{\ell} \left[G'(\zeta_{\ell}) - G(\zeta_{\ell}) \right] e^{\zeta_{\ell} 0^+} \quad (175)$$

where G' is the Green's function in the presence of the impurity and the trace is over the appropriate single-particle quantum numbers. Even for

the translationally invariant system G' will not be diagonal in the momenta because of the presence of the impurity. We may introduce the self-energy operator in the presence of the impurity according to

$$G'(\zeta_\ell) = \frac{1}{\zeta_\ell - h - \Sigma'(\zeta_\ell)} \quad (176)$$

where h is the single-particle Hamiltonian. It is convenient to separate out from Σ' the part which is independent of the impurities. We may then write

$$\Sigma'(\zeta_\ell) = \Sigma(\zeta_\ell) + \tilde{v}(\zeta_\ell) \quad (177)$$

Some graphs that contribute to $\tilde{v}(\zeta_\ell)$ are illustrated in Fig. 13. The interaction lines terminating at crosses indicate scatterings from the impurity. We note that just as for Σ one can show that $\Sigma'(\mu)$ is real. Therefore

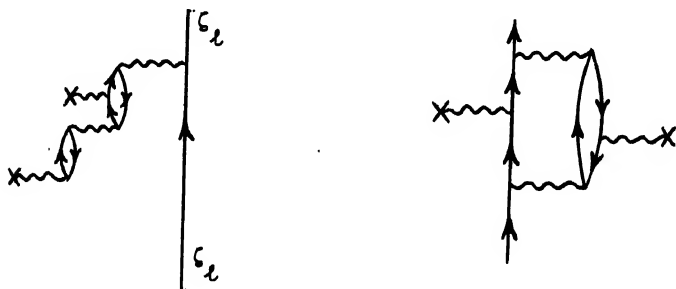


Fig. 13

$\tilde{v}(\mu)$ is real. One can simplify (175) by using the formula (167). Both in the presence and absence of the impurity the term involving $G(\partial/\partial\zeta)\Sigma$ is zero at $T = 0$. Then, using the same technique as in (169) to do the sum over ζ_ℓ one finds

$$\Delta N = \frac{1}{2\pi i} \text{Tr} \ln \left\{ 1 - 2\pi i \delta(h + \Sigma(\mu) - \mu) \tilde{v}(\mu) \cdot \left[1 - \frac{1}{\mu - h - \Sigma(\mu) + i\eta} \tilde{v}(\mu) \right]^{-1} \right\} \quad (178)$$

If we let φ_a represent the eigenstates of the model Hamiltonian $h + \Sigma(\mu)$, then the outgoing scattering states of the model Hamiltonian

$$h + \Sigma(\mu) + \tilde{v}(\mu) \quad (179)$$

with energy μ are given by

$$\psi_a^{(+)} = \varphi_a + \frac{1}{\mu - h - \Sigma(\mu) + i\eta} \tilde{v}(\mu) \psi_a^{(+)} \quad (180)$$

The S matrix for this scattering process is

$$S_{ba}(\mu) = \delta_{ba} - 2\pi i \delta(E_b - \mu) T_{ba} \quad (181)$$

where

$$T_{ba} = \left(\varphi_b, \tilde{v}(\mu) \psi_a^{(+)} \right) \quad (182)$$

From Eqs. (179) - (182) we see that (178) reduces to

$$\Delta N = \frac{1}{2\pi i} \text{Tr} \ln S(\mu) \quad (183)$$

In the case of spherical symmetry the eigenvalues of $S(\mu)$ are $\exp [2i\delta_\ell(k_F)]$ with multiplicity $2(2\ell+1)$ (the factor 2 from the sum over spin). Thus

$$\Delta N = \frac{2}{\pi} \sum_{\ell} (2\ell + 1) \delta_{\ell}(k_F) \quad (184)$$

Equation (184) is identical to Eq. (31). The Friedel sum rule (25) follows from (184) in the same way as it did from (31).

Thermodynamics

The thermodynamics of the system follows from the grand potential Ω which is defined according to

$$Z_G = \text{Tr} \left\{ e^{-\beta(H - \mu N)} \right\} = e^{-\beta\Omega(\mu, \beta, \text{Vol})} \quad (185)$$

In other words

$$\Omega = -\frac{1}{\beta} \ln Z_G \quad (186)$$

In order to generate a perturbation expansion for Ω it is convenient to use a well-known trick. We introduce an explicit coupling constant λ in the interaction term and write

$$H = H_0 + \lambda V \quad (187)$$

By differentiating (185) with respect to λ one finds

$$-\beta e^{-\beta\Omega} \frac{\partial\Omega}{\partial\lambda} = \text{Tr} \left\{ e^{-\beta(H-\mu N)} (-\beta V) \right\} \quad (188)$$

where the cyclic invariance of the trace has been used. From (188) one finds

$$\frac{\partial\Omega}{\partial\lambda} = \langle V \rangle \quad (189)$$

Introducing the notation $\Omega(\lambda=0) = \Omega_0$ one obtains from the integration of (189)

$$\Omega(\lambda) = \Omega_0 + \int_0^\lambda \frac{d\lambda'}{\lambda'} \langle \lambda' V \rangle \quad (190)$$

From the form of the two-body interaction,

$$\lambda V = \frac{1}{2} \int d^3\tau \int d^3\tau' v(\underline{x} - \underline{x}') \psi^\dagger(\underline{x}) \psi^\dagger(\underline{x}') \psi(\underline{x}') \psi(\underline{x}) \quad (191)$$

and the definition of the two-particle Green's function, we see that (190) may be written in the form[†]

$$\Omega = \Omega_0 - \frac{i}{2\tau} \int_0^\lambda \frac{d\lambda'}{\lambda'} v(\bar{1}, \bar{1}') G(\bar{1}, \bar{1}'; \bar{1}^+, \bar{1}'^+) \quad (192)$$

Now from the form (121) of the self-energy operator,

$$\Omega = \Omega_0 - \frac{i}{2\tau} \int_0^\lambda \frac{d\lambda'}{\lambda'} \Sigma(\bar{1}, \bar{1}') G(\bar{1}', \bar{1}'^+) \quad (193)$$

[†] The factor of τ in the denominator occurs because of the two time integrations in the integrand. One of these is collapsed by the definition (111) of $v(\bar{1}, \bar{1}')$ and the second gives a multiplicative factor of τ .

Because of the integration over the coupling constant, (193) is not a very useful form. The integration may be formally done by explicitly exhibiting all the dependence on the interaction. To do this we expand G according to (149) and define

$$\Sigma' = \Sigma + \Sigma G_0 \Sigma + \dots \quad (194)$$

All of the interaction lines are now contained in Σ' . We may break up Σ' according to

$$\Sigma' = \sum_n \Sigma'^{(n)} \quad (195)$$

where the superscript indicates the number of interaction lines. Then from (193) one has

$$\begin{aligned} \Omega &= \Omega_0 - \frac{i}{2\pi} \int_0^\lambda \frac{d\lambda'}{\lambda'} \Sigma'(\bar{1}, \bar{1}') G_0(\bar{1}', \bar{1}^+) \\ &= \Omega_0 - \frac{i}{2\pi} \sum_n \frac{1}{n} \Sigma'^{(n)}(\bar{1}, \bar{1}') G_0(\bar{1}', \bar{1}^+) \end{aligned} \quad (196)$$

Notice that the expansion of Ω consists of the contribution of vacuum to vacuum graphs (improper self-energy graphs closed on themselves) with, however, the extra factor $1/2n$ for an n th-order graph.

Equation (196) is still not a useful form for our purposes since it contains G_0 , whereas our aim is to express the thermodynamic properties in terms of the true spectrum described by G . It is possible to express Ω in terms of G as was first pointed out by Luttinger⁽³⁾ following the ideas of Lee and Yang.⁽¹¹⁾ To do this we first notice that one can think of Σ as the sum of the contribution of skeleton graphs (i.e., graphs in which the internal lines have no self-energy insertions) with each internal line replaced by the true Green's function. Then, consider the quantity

$$\Omega' = - \frac{i}{\pi} \sum_\nu \frac{1}{2\nu} \Sigma^{(\nu)}(\bar{1}, \bar{1}') G(\bar{1}', \bar{1}^+) \quad (197)$$

where ν is the number of explicit interactions in a skeleton graph for Σ . The graphs that occur in the expansion of Ω' are identical to those that occur in the expansion for Ω . However, the weight factors are not the same. For example the graph in Fig. 14(a) has a factor of $1/4$ in the expansion (196); however in the expansion (197) the graph occurs twice as indicated in Fig. 14(b) and has a weight factor $1/2$ from each of these occurrences.

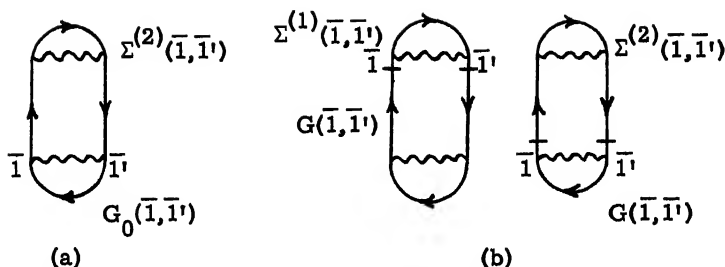


Fig. 14

Because of the overcounting, $\Omega' \neq \Omega - \Omega_0$. We shall see below that it is possible to compensate for the counting. Before doing this we note two properties of Ω' . First,

$$\begin{aligned} \frac{\delta \Omega'}{\delta G(\underline{p}, \zeta_\ell)} &= -\frac{i}{\tau} \sum_{\nu} \frac{2\nu}{2\nu} \Sigma^{(\nu)}(\underline{p}, \zeta_\ell) = -\frac{i}{\tau} \Sigma(\underline{p}, \zeta_\ell) \\ &= \frac{1}{\beta} \Sigma(\underline{p}, \zeta_\ell) \end{aligned} \quad (198)$$

Here we have used the facts that each term in $\Sigma^{(\nu)}$ has $(2\nu-1)$ electron lines and that

$$\sum_{\nu} \Sigma^{(\nu)} = \Sigma \quad (199)$$

Second,

$$\begin{aligned} \frac{\delta \Omega'}{\delta \Sigma(\underline{p}, \zeta_\ell)} &= \frac{\delta \Omega'}{\delta G(\underline{p}, \zeta_\ell)} \frac{\delta}{\delta \Sigma(\underline{p}, \zeta_\ell)} \frac{1}{G_0^{-1} - \Sigma(\underline{p}, \zeta_\ell)} \\ &= \frac{1}{\beta} \Sigma(\underline{p}, \zeta_\ell) [G(\underline{p}, \zeta_\ell)]^2 \end{aligned} \quad (200)$$

Now consider the quantity

$$Y \equiv \Omega' - \frac{1}{\beta} \sum_{\underline{p}, \ell} e^{\zeta_\ell 0^+} \left\{ \Sigma(\underline{p}, \zeta_\ell) G(\underline{p}, \zeta_\ell) + \ln [-G^{-1}(\underline{p}, \zeta_\ell)] \right\} \quad (201)$$

We shall show that $Y = \Omega$. First we note that

$$\frac{\delta Y}{\delta \Sigma(\underline{p}, \zeta_\ell)} = \frac{1}{\beta} \Sigma(\underline{p}, \zeta_\ell) [G(\underline{p}, \zeta_\ell)]^2 - \frac{1}{\beta} G(\underline{p}, \zeta_\ell)$$

$$-\frac{1}{\beta} \Sigma(\underline{p}, \zeta_l) \left[G(\underline{p}, \zeta_l) \right]^2 + \frac{1}{\beta} G(\underline{p}, \zeta_l) = 0 \quad (202)$$

Thus Y is stationary with respect to variations in Σ if G^{-1} is everywhere understood as $G_0^{-1} - \Sigma$. We see that the derivative of Y with respect to the coupling constant may be obtained by considering only those interactions that are not contained in Σ . Thus

$$\lambda \frac{\partial Y}{\partial \lambda} = \lambda \frac{\partial \Omega'}{\partial \lambda} \Big|_{\text{explicit}} \quad (203)$$

However from (197) and (196) we have

$$\lambda \frac{\partial \Omega'}{\partial \lambda} \Big|_{\text{explicit}} = \frac{1}{\beta} \Sigma(\bar{1}, \bar{1}') G(\bar{1}', \bar{1}^+) = \lambda \frac{\partial \Omega}{\partial \lambda} \quad (204)$$

Therefore if we can prove $Y(\lambda = 0) = \Omega(\lambda = 0)$ we have shown that $Y = \Omega$. Now from the definition (201)

$$Y(\lambda = 0) = -\frac{1}{\beta} \sum_{\underline{p}, \zeta_l} e^{\zeta_l 0^+} \ln(\epsilon_p - \zeta_l) \quad (205)$$

Doing the sum by the contour integral technique

$$Y(\lambda = 0) = -\frac{1}{2\pi i} \sum_{\underline{p}} \int_{\Gamma_0} d\zeta f^-(\zeta) \ln(\epsilon_p - \zeta) \quad (206)$$

Since

$$f^-(\zeta) = -\frac{1}{\beta} \frac{\partial}{\partial \zeta} \ln \left(1 + e^{-\beta(\zeta - \mu)} \right) \quad (207)$$

one obtains by a partial integration

$$\begin{aligned} Y(\lambda = 0) &= -\frac{1}{\beta} \frac{1}{2\pi i} \sum_{\underline{p}} \int_{\Gamma_0} d\zeta \frac{\zeta}{\zeta - \epsilon_p} \ln \left(1 + e^{-\beta(\zeta - \mu)} \right) \\ &= -\frac{1}{\beta} \sum_{\underline{p}} \ln \left(1 + e^{-\beta(\epsilon_p - \mu)} \right) \end{aligned} \quad (208)$$

But this is just the grand potential for a noninteracting system of fermions

$$\begin{aligned}\Omega_0 &= -\frac{1}{\beta} \ln Z_{G_0} = -\frac{1}{\beta} \ln \prod_{\underline{p}} \left(1 + e^{-\beta(\epsilon_{\underline{p}} - \mu)} \right) \\ &= -\frac{1}{\beta} \sum_{\underline{p}} \ln \left(1 + e^{-\beta(\epsilon_{\underline{p}} - \mu)} \right)\end{aligned}\quad (209)$$

The equality of Y and Ω is therefore proved.

Because of the stationarity property (202) the form (201) is very useful. We consider first an expansion of Ω in T at fixed μ and volume. Because of (202) we may replace Σ by Σ^0 (its value at $T = 0$) as far as the first temperature correction is concerned. Thus

$$\begin{aligned}\delta \Omega &= \delta \Omega' - \delta \left\{ \frac{1}{\beta} \sum_{\underline{p}, \ell} e^{\zeta_{\ell} 0^+} \right. \\ &\quad \times \left. \left[\Sigma^0(\underline{p}, \zeta_{\ell}) \frac{1}{\zeta_{\ell} - \epsilon_{\underline{p}} - \Sigma^0(\underline{p}, \zeta_{\ell})} + \ln(\epsilon_{\underline{p}} + \Sigma^0(\underline{p}, \zeta_{\ell}) - \zeta_{\ell}) \right] \right\}\end{aligned}\quad (210)$$

The first temperature correction of Ω' may be computed as follows. Since each line of a skeleton diagram contains a sum $(\beta^{-1} \cdot \Sigma_{\ell})$, the total first temperature correction to Ω' is obtained by correcting to first order the contribution of a single line, evaluating the rest of the diagram at $T = 0$, and finally summing over every line. This gives to first order

$$\delta \Omega' = \delta \left\{ \frac{1}{\beta} \sum_{\underline{p}, \ell} \frac{1}{\zeta_{\ell} - \epsilon_{\underline{p}} - \Sigma^0(\underline{p}, \zeta_{\ell})} \Sigma^0(\underline{p}, \zeta_{\ell}) \right\}\quad (211)$$

from (211) and (210) we find

$$\delta \Omega = -\delta \left\{ \frac{1}{\beta} \sum_{\underline{p}, \ell} e^{\zeta_{\ell} 0^+} \ln(\epsilon_{\underline{p}} + \Sigma^0(\underline{p}, \zeta_{\ell}) - \zeta_{\ell}) \right\}\quad (212)$$

The first temperature correction of the right-hand side is easily obtained by using the contour integral technique. We have

$$\begin{aligned}& -\frac{1}{\beta} \sum_{\underline{p}, \ell} e^{\zeta_{\ell} 0^+} \ln(\epsilon_{\underline{p}} + \Sigma^0(\underline{p}, \zeta_{\ell}) - \zeta_{\ell}) \\ &= -\frac{1}{2\pi i} \sum_{\underline{p}} \int_{\Gamma_0} d\zeta f^-(\zeta) e^{\zeta 0^+} \ln(\epsilon_{\underline{p}} + \Sigma^0(\underline{p}, \zeta) - \zeta)\end{aligned}$$

$$= -\frac{1}{2\pi i} \sum_{\underline{p}} \int_{-\infty}^{\infty} dx \ln \left[\frac{\epsilon_{\underline{p}} + \Delta_{\underline{p}}(x) + \frac{i}{2} \Gamma_{\underline{p}}(x) - x}{\epsilon_{\underline{p}} + \Delta_{\underline{p}}(x) - \frac{i}{2} \Gamma_{\underline{p}}(x) - x} \right] \frac{1}{e^{\beta(x-\mu)} + 1} \quad (213)$$

From the Sommerfield expansion

$$\frac{1}{e^{\beta(x-\mu)} + 1} = \theta(\mu - x) - \frac{\pi^2 k_B^2 T^2}{6} \frac{\partial}{\partial x} \delta(\mu - x) + \dots \quad (214)$$

one finds the first temperature correction to (213). Thus we see that

$$\Omega(T, \mu, \text{Vol}) = \Omega(T=0, \mu, \text{Vol}) - \frac{\pi^2 k_B^2 T^2}{6} \sum_{\underline{p}} \delta(\mu - E_{\underline{p}}) + O(T^4) \quad (215)$$

Let us contrast this behavior with that of the noninteracting system. An expansion similar to the above for the expression (205) leads to

$$\begin{aligned} \Omega_0(T, \mu, \text{Vol}) &= \sum_{\underline{p}} (\epsilon_{\underline{p}} - \mu) \theta(\mu - \epsilon_{\underline{p}}) - \frac{\pi^2 k_B^2 T^2}{6} \sum_{\underline{p}} \delta(\mu - \epsilon_{\underline{p}}) \\ &+ O(T^4) \end{aligned} \quad (216)$$

The first term of (215) has no simple structure analogous to that of the first term of (216). However the second terms have essentially the same form. This means that the entropy for the interacting system has the same form (to lowest order in the temperature) as that for the noninteracting system, for the entropy is given by

$$S(\mu, T, \text{Vol}) = - \left(\frac{\partial \Omega}{\partial T} \right)_{\mu, \text{Vol}} = \frac{\pi^2 k_B^2 T}{3} \sum_{\underline{p}} \delta(\mu - E_{\underline{p}}) + O(T^3) \quad (217)$$

Note that this may be written in the form

$$S(\mu, T, \text{Vol}) = -k_B \sum_{\underline{p}} \left[f^-(E_{\underline{p}}) \ln f^-(E_{\underline{p}}) + f^+(E_{\underline{p}}) \ln f^+(E_{\underline{p}}) \right] + O(T^3) \quad (218)$$

where

$$f^{\pm}(E_{\underline{p}}) = \frac{1}{e^{\mp\beta(E_{\underline{p}} - \mu)} + 1} \quad (219)$$

We here make explicit contact with the Landau theory of Sec. 1, and we see that E_p , determined by (155), plays the role of the quasi-particle dispersion curve at $T = 0$. Notice, however, that we have only proved (218) to lowest order in T . In the Landau theory the single-particle energies are temperature dependent and an expression of the form (218) is assumed to hold for the next temperature correction as well. Bloch and de Dominicis⁽¹²⁾ have shown that it is possible to write the entropy in the many-particle theory in the form (218) for all temperatures. In their formalism, however, there is no identification of E_p as a single-particle energy and some of their generality must therefore be regarded as purely formal. It is presumably possible to include the T^3 term in the form (218) by identifying E_p as being given by the pole of the Green's function at small but finite temperatures. We have not, however, looked into this possibility, and here will confine ourselves to the lowest temperature correction.

From (215) the rest of the thermodynamics of the Landau theory also follows. For the specific heat we get

$$C = T \left(\frac{\partial S}{\partial T} \right)_{N, \text{Vol}} = \frac{\pi^2 k_B^2}{3} T \sum_{\underline{p}} \delta(\mu^0 - E_{\underline{p}}) + O(T^3) \quad (220)$$

where μ^0 is the chemical potential for the interacting system at $T = 0$ with a fixed number of particles.

The internal energy may be calculated from

$$dE|_{N, \text{Vol}} = T dS|_{N, \text{Vol}} \quad (221)$$

using (217) we obtain

$$E(T, N, Vol) = E(T = 0, N, Vol) + \frac{\pi^2 k_B^2 T^2}{6} \times \sum \delta(E_{\underline{p}} - \mu^0) + O(T^4) \quad (222)$$

The change in energy at constant chemical potential can be obtained from the identities

$$\left(\frac{\partial E}{\partial T}\right)_{\mu} = \left(\frac{\partial E}{\partial T}\right)_N + \mu \left(\frac{\partial N}{\partial T}\right)_{\mu} \quad (223)$$

$$\left(\frac{\partial N}{\partial T}\right)_{\mu} = - \frac{\partial^2 \Omega}{\partial T \partial \mu} = \frac{\pi^2 k_B^2 T}{3} \sum_{\underline{p}} \frac{\partial}{\partial \mu} \delta(\mu - E_{\underline{p}}) \quad (224)$$

From (222), (223), and (224) one finds

$$\delta E|_{T, \mu, Vol} = \sum_{\underline{p}} E_{\underline{p}} \delta f^-(E_{\underline{p}}) + O(T^4) \quad (225)$$

which is again an equation of the Landau theory. For the number of particles we have using (171) and (224)

$$N(T, \mu, Vol) = \sum_{\underline{p}} f^-(E_{\underline{p}}) + O(T^4) \quad (226)$$

The expression (18) for the compressibility at low temperatures follows from (226) as in Sec. 1.

4. TRANSPORT PROPERTIES

We have seen that we can at least formally identify a certain energy spectrum in the neighborhood of a surface in momentum space such that

$$\delta E = \sum_{\underline{p}} E_{\underline{p}} \delta f^-(E_{\underline{p}}) + O(T^4) \quad (225)$$

where the single-particle-like energies $E_{\underline{p}}$ in principle are determined from the real part of the self-energy operator:

$$E_{\underline{p}} = \epsilon_{\underline{p}} + \Delta_{\underline{p}}(E_{\underline{p}}) \quad (227)$$

It was necessary to assume that Eq. (227) had a solution and that this solution was unique. Once this was done we could, making use of the structure of perturbation theory, express the thermodynamical properties of a system at low temperatures in terms of the weak excitations $E_{\underline{p}}$. In the process of doing this we have been able to substantiate the remarks made in the Introduction about why it is that the usual theory of metals works. It may be worth while to repeat those remarks here. The ground state of the ideal noninteracting system has a very simple structure. For the physical interacting system, on the other hand, the ground state is highly correlated, extremely complicated and essentially not known. But no particular use is made of the ground state, the only properties that are needed are the energies and numbers of low-lying excited states, and it follows from the assumptions we made that there is a one-to-one correspondence between the low-lying excited states of the ideal system and the interacting system.

We shall now discuss the transport properties of our interacting system. In particular we shall make a careful calculation of the residual resistivity. Our goal here will also be to express the transport properties in terms of the low-lying excited states. The customary treatment of transport properties makes use of the Boltzmann equation. We shall take a more fundamental starting point. The basic ideas of the method we shall use go back to the Kramers-Heisenberg semiclassical theory of dispersion. The use of such methods for the calculation of transport coefficients has been more recently advocated by Kubo.⁽¹³⁾

The conductivity is given by the response to an external electric field. We write for the interaction term.

$$\mathcal{H}(t) = \frac{1}{2m} \sum_i \left(\underline{p}_i - \frac{e}{c} \underline{A}(\underline{x}_i, t) \right)^2 - \sum_i \frac{p_i^2}{2m} \quad (228)$$

where we choose a gauge so that the scalar potential is zero and

$$\underline{E} = -\frac{1}{c} \frac{\partial \underline{A}}{\partial t} \quad (229)$$

$$\underline{A}(\underline{x}, -\infty) = 0 \quad (230)$$

In a system with Hamiltonian (49) there is no way of dissipating momentum. The d-c conductivity is therefore imaginary and infinite, corresponding to an acceleration of the electrons in the system. We shall be interested in the finite d-c conductivity which results when a random distribution of impurities is added to the system described by Eq. (2).

We separate the current into two parts. In the first term the external field does not occur explicitly

$$\underline{j}(\underline{x}) = \frac{e}{2mi} \left[\psi^+(\underline{x}) \nabla \psi(\underline{x}) - (\nabla \psi^+(\underline{x})) \psi(\underline{x}) \right] \quad (231)$$

and the total current operator is given by

$$\begin{aligned} \underline{J}(\underline{x}; t) &= \frac{e}{m} \sum_i \left(\underline{p}_i - \frac{e}{c} \underline{A}(\underline{x}, t) \right) \delta(\underline{x} - \underline{x}_i) \\ &= \underline{j}(\underline{x}) - \frac{e^2}{mc} \psi^+(\underline{x}) \psi(\underline{x}) \underline{A}(\underline{x}, t) \end{aligned} \quad (232)$$

The measured current is

$$\langle \underline{J}(\underline{x}, t) \rangle = \text{Tr} \left\{ \rho(t) \underline{J}(\underline{x}, t) \right\} \quad (233)$$

where the density matrix satisfies the Liouville equation

$$i \frac{\partial \rho}{\partial t} = [H + \mathcal{H}(t), \rho] \quad (234)$$

It is convenient to use the "interaction representation"

$$\rho_I(t) = e^{iHt} \rho(t) e^{-iHt} \quad (235)$$

Simple algebra shows that

$$i \frac{\partial \rho_I}{\partial t} = [\mathcal{H}_I(t), \rho_I(t)] \quad (236)$$

where

$$\mathcal{H}_I(t) = e^{iHt} \mathcal{H}(t) e^{-iHt} \quad (237)$$

We wish to calculate ρ only up to terms linear in \mathcal{H} :

$$\rho_I(t) = \rho_0 - i \int_{-\infty}^t dt' [\mathcal{H}_I(t'), \rho_0] + O(\mathcal{H}^2) \quad (238)$$

Similarly up to terms linear in the external field

$$\langle \underline{J}(\underline{x}, t) \rangle = -i \int_{-\infty}^t dt' \langle [\underline{j}_I(\underline{x}, t), \mathcal{H}_I(t')] \rangle = \rho_0 - \frac{ne^2}{mc} \underline{A}(\underline{x}, t) \quad (239)$$

Here the expectation values are to be taken with respect to the unperturbed density matrix, and only the term in \mathcal{H}_I that is linear in the external field is to be taken into account. Furthermore n is the particle number density (uniform in the unperturbed system). With

$$\underline{A} = -\frac{c}{i\omega} \underline{E} e^{i\omega t} \quad (240)$$

and writing

$$\langle \underline{J}(\underline{q}, \omega) \rangle = \underline{\sigma}(\underline{q}, \omega) \cdot \underline{E}(\underline{q}, \omega) \quad (241)$$

we see that

$$\begin{aligned} \sigma_{\mu\nu}(\underline{q}, \omega) &= i \int_{-\infty}^0 \langle [j_{\mu}(\underline{q}, 0), j_{\nu}(-\underline{q}, t)] \rangle \rho_0 \\ &\times \frac{e^{i\omega t}}{i\omega} dt - \frac{ne^2}{mc} \frac{c}{i\omega} \delta_{\mu\nu} \end{aligned} \quad (242)$$

Care must be taken in extracting the ordinary d-c conductivity σ from the tensor of Eq. (242). To see where the difficulty lies consider the perturbing electric field, Eq. (240). This may be broken up into its solenoidal and irrotational parts \underline{E}_{\perp} and $\underline{E}_{\parallel}$ respectively ($\nabla \cdot \underline{E}_{\perp} = 0$, $\nabla \times \underline{E}_{\parallel} = 0$).

Since the Coulomb interactions between electrons are included in the basic Hamiltonian (49), $\underline{E}_{\parallel}$ is the electric field due to external charges whereas \underline{E}_{\perp} is the total transverse field in the medium. Now the tensor (242)

contains terms that have the effect of screening $\underline{E}_{\parallel}$. Since the usual conductivity is the coefficient of proportionality between the current density

and the total or screened electric field, the screening terms must be omitted. The tensor structure of these terms is $q_{\mu} q_{\nu} G(q)/q^2$. They arise from the long wavelength part of the Coulomb interaction. One can simply avoid these terms altogether by setting $q = 0$. This is because the over-all charge neutrality of the system requires that $v(q = 0) = 0$. (See the footnote on page 353.) By using this device and introducing

$$\underline{j}(t) = \int d^3x \underline{j}_I(\underline{x}, t) = \underline{j}(q = 0, t) \quad (243)$$

we have as the formal expression for the long wavelength conductivity

$$\sigma(\omega) = \frac{i}{3V\Omega I} \int_{-\infty}^0 < [\underline{j}(0) \cdot \underline{j}(t)] > \frac{e^{-i\omega t}}{i\omega} - \frac{ne^2}{mc} \frac{c}{i\omega} \quad (244)$$

(The factor $1/3$ comes from an average over space directions.) In the d-c limit $\omega \rightarrow 0$, and the last term in (244) diverges as $1/\omega$. If we expand the exponential in powers of ωt , the leading term will also diverge as $1/\omega$. Since the system we are dealing with certainly has a finite d-c conductivity, the two divergences must cancel. (We explicitly show this in Appendix C.) Keeping the finite part in the $\omega \rightarrow 0$ limit, we have

$$\begin{aligned} \sigma(0) &= \frac{2}{3V\Omega I} \frac{1}{2i} \int_{-\infty}^0 t \left\{ < \underline{j}(0) \cdot \underline{j}(t) > - \text{c.c.} \right\} dt \\ &= \frac{2}{3V\Omega I} \text{Im} \int_{-\infty}^0 t < \underline{j}(0) \cdot \underline{j}(t) > dt \end{aligned} \quad (245)$$

The expectation value in (245) is a special case of the two-particle Green's function. It is convenient to introduce a function

$$\begin{aligned} F(t) &= T < \underline{j}(0) \cdot \underline{j}(t) > \\ &\times \begin{cases} < \underline{j}(t) \cdot \underline{j}(0) > = F_{>}(t) & t > 0 \\ < \underline{j}(0) \cdot \underline{j}(t) > = F_{<}(t) & t < 0 \end{cases} \end{aligned} \quad (246)$$

It is easily seen that

$$F_{>}^*(t) = F_{<}(t) = F_{>}(-t) \quad (247)$$

As a result

$$F_{\geq}(\omega) = F_{\leq}(-\omega) = F_{\geq}^*(\omega) \quad (248)$$

We get

$$\begin{aligned}
 \sigma(\omega \rightarrow 0) &= \frac{2}{3\text{Vol}} \text{Im} \int_{-\infty}^0 dt \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} t F_{<}(\omega) e^{-i\omega t} \\
 &= \frac{2\pi}{3\text{Vol}} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} F_{<}(\omega) \frac{\partial}{\partial \omega} \delta(\omega) \\
 &= \lim_{\omega \rightarrow 0} \frac{1}{6\text{Vol}} \frac{F_{>}(\omega) - F_{<}(\omega)}{\omega}
 \end{aligned} \tag{249}$$

where in the last two steps we have used (248). It is useful to extend to complex time differences.[†] With $\tau = -i\beta$ we get the periodicity condition

$$F_{>}(t + \tau) = F_{>}(t) \tag{250}$$

and

$$F_{>}(\omega) = e^{-\omega \tau} F_{<}(\omega) \tag{251}$$

Introducing the spectral weight function

$$F_{>}(\omega) = \frac{\varphi(\omega)}{1 - e^{-i\omega \tau}} \tag{252}$$

Eq. (249) becomes simply

$$\sigma(0) = \frac{1}{6\text{Vol}} \lim_{\omega \rightarrow 0} \frac{\varphi(\omega)}{\omega} \tag{253}$$

It will prove convenient to calculate the Fourier coefficients in the restricted time region $-\tau < t < \tau$, where as before the inequality signs refer to the negative imaginary part of t .

[†] Using the same criteria as for the single-particle Green's function, $F_{>}$ is continued into the lower half-plane whereas $F_{<}$ is continued into the upper half-plane.

$$F(t) = \frac{1}{\tau} \sum_{\mathbf{m}} e^{-i\nu_{\mathbf{m}} t} F(\nu_{\mathbf{m}}) \quad (254)$$

where

$$\nu_{\mathbf{m}} = \frac{2\pi\mathbf{m}}{\tau} \quad \mathbf{m} = 0, \pm 1, \pm 2, \dots \quad (255)$$

The inverted relation is

$$F(\nu_{\mathbf{m}}) = \int_0^{\tau} F_{>}(t) e^{i\nu_{\mathbf{m}} t} dt \quad (256)$$

Simple algebra gives

$$F(\nu_{\mathbf{m}}) = i \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{\varphi(\omega)}{\nu_{\mathbf{m}} - \omega} \quad (257)$$

Continuing to the ν plane

$$F(\nu) = i \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{\varphi(\omega)}{\nu - \omega} \quad (258)$$

and taking the discontinuity across the cut

$$F(\lambda + i\eta) - F(\lambda - i\eta) = \varphi(\lambda) \quad (259)$$

so that φ can be directly obtained from $F(t)$ with t on the imaginary axis.

An Approximate Calculation of Residual Resistivity

So far the discussion has been quite general; what we want to do is calculate the impurity resistance. Therefore we specialize to a Hamiltonian of the form

$$H = H_0 + V + V_{\text{imp}} \quad (260)$$

$$\begin{aligned}
 V_{\text{imp}} &= \sum_i \sum_{n=1}^{N_i} u(\underline{x}_i - \underline{R}_n) = \int \sum_{n=1}^{N_i} \psi^\dagger(\underline{x}) u(\underline{x} - \underline{R}_n) \psi(\underline{x}) d^3x \\
 &= \sum_{n=1}^{N_i} \sum_{\underline{k}, \underline{q}} a_{\underline{k}+\underline{q}}^+ a_{\underline{k}} u(\underline{q}) e^{-i\underline{q} \cdot \underline{R}_n}
 \end{aligned} \tag{261}$$

In Eq. (261) we have considered a system with a fixed distribution of impurities. For such a system the current that one would calculate from the relation (239) would be the microscopic current that varies over atomic dimensions. Now for the d-c case we are interested only in the average of this microscopic current over the whole system. One can imagine the system divided into many small but macroscopic regions. The average current we want can then be thought of as an average of the average current in each of the small regions. But this double average is just an ensemble average over systems with different distributions of impurities. Let us see how such an ensemble average may be performed.

Consider first the case of particles which do not interact with each other. The contributions to the single-particle Green's function before averaging are pictured in Fig. 15. In first order we have to average over the impurity position a term

$$\langle u(\underline{q}) e^{-i\underline{q} \cdot \underline{R}_n} \rangle_{\text{imp}} \tag{262}$$

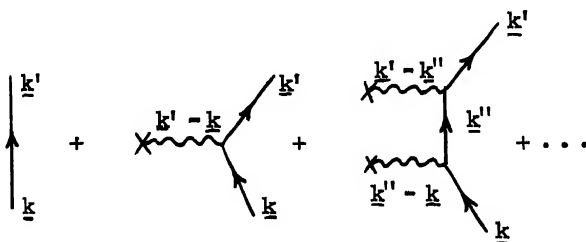


Fig. 15

There are N_i such terms and the average over position is performed by an integral $(1/\text{Vol}) \int d^3R$, i.e., we get

$$\langle \sum_n u(\underline{q}) e^{-i\underline{q} \cdot \underline{R}_n} \rangle_{\text{imp}} = \frac{N_i}{\text{Vol}} \int d^3R u(\underline{q}) e^{-i\underline{q} \cdot \underline{R}}$$

$$= (2\pi)^3 n_i u(0) \delta(\underline{q}) \quad (263)$$

where $n_i = N_i/\text{Vol}$. This term leads to a constant shift in the zero of energy and can simply be removed by a redefinition of the chemical potential. The second-order term will include an average

$$\left\langle \sum_{n,m} u(\underline{q}) u(\underline{k}) e^{-i\underline{q} \cdot \underline{R}_n} e^{-i\underline{k} \cdot \underline{R}_m} \right\rangle_{\text{imp}} \quad (264)$$

In the sum over impurities in (264) there are two kinds of terms, those with $n = m$ and with $n \neq m$. The first kind gives

$$\begin{aligned} n_i \int d^3 R u(\underline{q}) u(\underline{k}) e^{-i(\underline{q}+\underline{k}) \cdot \underline{R}} \\ = n_i u(\underline{q}) u(\underline{k}) \delta(\underline{q} + \underline{k}) (2\pi)^3 \end{aligned} \quad (265)$$

In the case $n \neq m$ averaging gives instead

$$\frac{N_i(N_i - 1)}{\text{Vol}^2} \int d^3 R_1 d^3 R_2 u(\underline{q}) u(\underline{k}) e^{-i\underline{q} \cdot \underline{R}_1} e^{-i\underline{k} \cdot \underline{R}_2} \quad (266)$$

Equation (266) is of $O(n_i^2)$ and will be neglected. It is convenient to draw graphs corresponding to the impurity averaging. In general we get terms corresponding to one, two, three, etc. scatterings from the same impurity and then also terms giving scattering from more than one impurity. In the limit we are considering (low impurity density) only the first type of terms is important. Graphically the picture is given in Fig. 16. The result of the averaging is that the sum over all momentum transfers in any graph is zero.

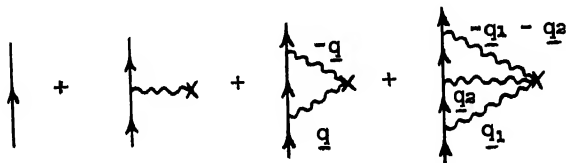


Fig. 16

The contribution to the self-energy operator of the graph with one interaction line is just a trivial shift in the chemical potential. For the two-interaction graph we have

$$\Sigma^{(2)}(\underline{k}, \zeta_{\ell}) = n_i \int \frac{d^3 q}{(2\pi)^3} \frac{|u(q)|^2}{\zeta_{\ell} - \epsilon_{\underline{k}-\underline{q}}} \quad (267)$$

which we formally write with $|\underline{k}\rangle$ an eigenstate of the single-particle Hamiltonian h

$$n_i < \underline{k} | u \frac{1}{\zeta_{\ell} - h} u | \underline{k} > \quad (268)$$

The discontinuity across the branch cut in (267) is given by $i\Gamma_{\underline{k}}^{(u)}$, where

$$\Gamma_{\underline{k}}^{(u)} = 2\pi n_i \int \frac{d^3 p}{(2\pi)^3} |u(p - \underline{k})|^2 \delta(\omega - \epsilon_p) \quad (269)$$

This is just n_i times the scattering rate for one impurity in the Born approximation. The totality of graphs in Fig. 16 leads to an exact treatment of the scattering by one impurity. To see this we write down the contributions to the self-energy of the graphs of Fig. 16

$$\begin{aligned} \Sigma(\underline{k}, \zeta_{\ell}) = & n_i < \underline{k} | u | \underline{k} > + n_i < \underline{k} | u \frac{1}{\zeta_{\ell} - h} u | \underline{k} > \\ & + n_i < \underline{k} | u \frac{1}{\zeta_{\ell} - h} u \frac{1}{\zeta_{\ell} - h} u | \underline{k} > + \dots \end{aligned} \quad (270)$$

We see that we can make a formal analogy with the usual theory of scattering. Just above the real axis we have

$$\Sigma(\underline{k}, \omega + i\epsilon) = n_i < \underline{k} | u | \underline{k}^{(+)} > \quad (271)$$

where

$$|\underline{k}^{(+)} > = |\underline{k} > + \frac{1}{\omega - h + i\epsilon} | \underline{k}^{(+)} > \quad (272)$$

Therefore, if we consider the t matrix for scattering with the model Hamiltonian $h + u$

$$\Sigma(\underline{k}, \omega + i\epsilon) = n_1 t_{\underline{k}\underline{k}}^{(+)}(\omega) \quad (273)$$

and

$$\Gamma_{\underline{k}}(\omega) = 2 \operatorname{Im} \Sigma(\underline{k}, \omega + i\epsilon) = -2n_1 \operatorname{Im} t_{\underline{k}\underline{k}}^{(+)}(\omega) \quad (274)$$

Now, the well-known optical theorem states

$$\operatorname{Im} t_{\underline{k}\underline{k}}(\omega) = -\pi \sum_{\underline{k}'} |t_{\underline{k}\underline{k}'}|^2 \delta(\omega - \epsilon_{\underline{k}'}) \quad (275)$$

so that

$$\Gamma_{\underline{k}}(\omega) = 2\pi n_1 \int \frac{d^3 k'}{(2\pi)^3} |t_{\underline{k}\underline{k}'}|^2 \delta(\omega - \epsilon_{\underline{k}'}) \quad (276)$$

Before proceeding, we give a direct derivation of the optical theorem (275). The reason for doing this is that what we shall thereby learn will be useful later.

Define

$$t_{\underline{k}'\underline{k}}(\zeta) = \langle \underline{k}' | u | \underline{k} \rangle + \langle \underline{k}' | u \frac{1}{\zeta - \epsilon_{\underline{k}} - u} u | \underline{k} \rangle + \dots \quad (277)$$

The diagonal term $t_{\underline{k}\underline{k}}$ is real when ζ is real; therefore

$$\operatorname{Im} t_{\underline{k}\underline{k}}(\omega + i\epsilon) = \frac{1}{2i} \{ t_{\underline{k}\underline{k}}(\omega + i\epsilon) - t_{\underline{k}\underline{k}}(\omega - i\epsilon) \} \quad (278)$$

The computation of (278) for the case with only two interaction lines [Σ given by (267)] is trivial since there is only one energy dominator, and the result is given in (269). Consider now the more complicated case with three lines. What we have to compute is the discontinuity

$$\frac{1}{2i} \left\{ \frac{1}{\omega - \epsilon_{\underline{k}+\underline{q}_1} + i\epsilon_1} \cdot \frac{1}{\omega - \epsilon_{\underline{k}+\underline{q}_1+\underline{q}_2} + i\epsilon_2} - \frac{1}{\omega - \epsilon_{\underline{k}+\underline{q}_1} - i\epsilon_1} \cdot \frac{1}{\omega - \epsilon_{\underline{k}+\underline{q}_1+\underline{q}_2} - i\epsilon_2} \right\} \quad (279)$$

Add and subtract

$$\frac{1}{2i} \left\{ \frac{1}{\omega - \epsilon_{\underline{k}+\underline{q}_1} + i\epsilon_1} - \frac{1}{\omega - \epsilon_{\underline{k}+\underline{q}_1+\underline{q}_2} - i\epsilon_2} \right\} \quad (280)$$

to give

$$\frac{\pi}{\omega - \epsilon_{\underline{k}+\underline{q}_1} + i\epsilon_1} \delta(\omega - \epsilon_{\underline{k}+\underline{q}_1+\underline{q}_2}) + \frac{\pi}{\omega - \epsilon_{\underline{k}+\underline{q}_1+\underline{q}_2} - i\epsilon_1} \delta(\omega - \epsilon_{\underline{k}+\underline{q}_1}) \quad (281)$$

In general one can calculate the discontinuity across the real axis (278) of terms containing many overlapping branch cuts by first displacing the branch cuts by a small amount and then crossing them one by one. The third-order case discussed above shows how this may be done. It is straightforward to apply this method to each of the terms of (277). Then one finds

$$\begin{aligned} \text{Im } t_{\underline{k}\underline{k}}^{(w)} &= -n \sum_{\underline{p}} t_{\underline{k}\underline{p}}^{(w+i\epsilon)} \delta(\omega - \epsilon_{\underline{p}}) t_{\underline{p}\underline{k}}^{(w-i\epsilon)} \\ &= -\pi \sum_{\underline{p}} |t_{\underline{k}\underline{p}}^{(w)}|^2 \delta(\omega - \epsilon_{\underline{p}}) \end{aligned} \quad (282)$$

which is just the optical theorem.

When going over to the interacting system we can borrow the results from our discussion of the Friedel sum rule. With $|\underline{k}\rangle$ an eigenstate of the model Hamiltonian $h + \Delta(\mu)$ and

$$t_{\underline{k}\underline{k}'} = \langle \underline{k} | \tilde{v}(\mu) | \underline{k}' \rangle \quad (283)$$

we get

$$\Gamma_{\underline{k}}(\mu) = 2\pi n_i \int \frac{d^3 p}{(2\pi)^3} |t_{\underline{k}p}(\mu)|^2 \delta(\mu - \epsilon_p - \Delta_p(\mu)) \quad (284)$$

For our further discussion we need an estimate of the order of magnitude of $\Gamma_{\underline{k}}(\mu)$ and its \underline{k} dependence. In the Thomas-Fermi approximation one substitutes for $t_{\underline{k}\underline{k}'}(\mu)$ the screened potential

$$t_{\underline{k}\underline{k}'}(\mu) \approx \frac{4\pi e^2 Z}{(\underline{k} - \underline{k}')^2 + \kappa^2} \quad (285)$$

where κ is the inverse Thomas-Fermi screening length. For our purposes the only thing we have to know about κ is that it is a quantity of the order of magnitude k_F . Substituting (285) into (284) and performing the integration.

$$\Gamma_{\underline{k}}(\mu) \approx 8\pi m n_i e^4 Z \left\{ \frac{4k_F}{[(k - k_F)^2 + \kappa^2] [(k + k_F)^2 + \kappa^2]} \right\} \quad (286)$$

From (286) we see that $\Gamma_{\underline{k}}(\mu)$ is peaked at $k = k_F$ with a maximum of the order of magnitude. Furthermore the function has a width of the order k_F . This means that $\Gamma_{\underline{k}}(\mu)$ is a very smoothly varying function of the momentum. In particular we can put $\Gamma_{\underline{k}}(\mu)$ equal to a constant over a range in the corresponding energy of the order of magnitude of $\Gamma_{\underline{k}}(\mu)$ when integrating over momenta close to the Fermi surface.

Fermi surface. Knowing roughly the properties of $\Gamma_{\underline{k}}(\mu)$ we can give a simple physical interpretation of it in terms of a mean free path of quasi particles, and it will turn out that the mean free path is just the important physical quantity for determining the d-c conductivity. We have shown that

$$G(\underline{k}, \mu + i\eta) = \frac{1}{\mu - \epsilon_{\underline{k}} - \Delta_{\underline{k}}(\mu) + \frac{i}{2} \Gamma_{\underline{k}}(\mu)} \quad (287)$$

In order to give a physical interpretation it is convenient to work with the spectral weight function

$$A(\underline{k}, \mu) = \frac{\Gamma_{\underline{k}}(\mu)}{\left(\mu - \epsilon_{\underline{k}} - \Delta_{\underline{k}}(\mu)\right)^2 + \frac{\Gamma_{\underline{k}}^2}{4}} \quad (288)$$

In configuration space

$$A(\underline{x}, \mu) = \int \frac{d^3 k}{(2\pi)^3} \frac{\Gamma_{\underline{k}}(\mu)}{\left(\mu - \epsilon_{\underline{k}} - \Delta_{\underline{k}}(\mu)\right)^2 + \frac{\Gamma_{\underline{k}}^2}{4}} e^{i\underline{k} \cdot \underline{x}} \quad (289)$$

The angular integrations are easily performed:

$$A(\underline{x}, \mu) = \frac{4\pi}{(2\pi)^3} \int d\epsilon_{\underline{k}} \frac{dk}{d\epsilon_{\underline{k}}} k^2 \frac{e^{ikx} - e^{-ikx}}{2ik} \frac{1}{i} \times \left[\frac{1}{\mu - \epsilon_{\underline{k}} - i\frac{\Gamma_{\underline{k}}}{2}} - \frac{1}{\mu - \epsilon_{\underline{k}} + i\frac{\Gamma_{\underline{k}}}{2}} \right] \quad (290)$$

where we have substituted

$$\epsilon_{\underline{k}} \equiv \epsilon_{\underline{k}} + \Delta_{\underline{k}}(\mu) \quad (291)$$

and

$$dk = \frac{dk}{d\epsilon_{\underline{k}}} d\epsilon_{\underline{k}} \quad (292)$$

The integrals (290) can be closed in the upper and lower half-planes, respectively for the terms containing e^{ikx} and e^{-ikx} . The value of (290) becomes just the sum of the residues at the poles

$$\epsilon_k = \mu \pm i \frac{\Gamma_k}{2} \quad (293)$$

We have

$$k\left(\mu \pm i \frac{\Gamma_k}{2}\right) = k_F \pm i \frac{\Gamma_k}{2} \frac{\partial k}{\partial \epsilon_k} \bigg|_{\epsilon_k = \mu} \quad (294)$$

Remembering that the quasi-energies are given not by ϵ_k but by

$$E_k = \epsilon_k + \Delta_k(E_k) \approx \epsilon_k + (E_k - \mu) \frac{\partial \Delta_k(E_k)}{\partial E_k} \bigg|_{E_k = \mu} \quad (295)$$

we have

$$\begin{aligned} \frac{\partial E_k}{\partial k} \bigg|_{k=k_F} \left(1 - \frac{\partial \Delta_k(E_k)}{\partial E_k} \bigg|_{E_k = \mu} \right) \\ = \frac{1}{Z_{k_F}} \frac{\partial E_k}{\partial k} \bigg|_{k=k_F} = \frac{\partial \epsilon_k}{\partial k} \bigg|_{k=k_F} \end{aligned} \quad (296)$$

The group velocity of a wave packet of quasi-particles is

$$v_k = \frac{\partial E_k}{\partial k} \bigg|_{k=k_F} \approx Z_{k_F} \frac{\partial \epsilon_k}{\partial k} \bigg|_{k=k_F} \quad (297)$$

and Eq. (294) therefore becomes

$$k \left(\mu \pm i \frac{\Gamma_k}{2} \right) \approx k_F \pm \frac{i \Gamma_{k_F} Z_{k_F}}{2v_F} \quad (298)$$

All the relations from (294) to (298) are valid because $\Gamma_k(\mu)$ is a slowly varying function in momentum space and

$$\Gamma_k(\mu) \ll \mu \quad (299)$$

When taking the residue of (290) we can put $k = k_F$ everywhere except in the exponentials and we get with

$$\bar{\Gamma}_{k_F} \equiv Z_{k_F} \Gamma_{k_F}(\mu) \quad (300)$$

$$\begin{aligned} A_{\text{imp}}(\underline{x}, \mu) &= A(\underline{x}, \mu) e^{-x \bar{\Gamma}_{k_F} / 2v_F} \\ &= A(\underline{x}, \mu) e^{-x / 2\ell_F} \quad (x = |\underline{x}|) \end{aligned} \quad (301)$$

where $A(\underline{x}, \mu)$ is the spectral weight function and

$$\ell_F = \frac{v_F}{\bar{\Gamma}_{k_F}} \quad (302)$$

may be interpreted as the mean free path. To see this note that because of (301) we also must have

$$G_{\text{imp}}(\underline{x}, \mu) = G(\underline{x}, \mu) e^{-x/2\ell_F} \quad (303)$$

We interpreted the single-particle Green's function as a probability amplitude for a single-particle excitation inserted at the origin to be found later

at the point \underline{x} . From this interpretation it is seen that l_F is just the mean free path of a quasi particle with energy μ . (The factor $1/2$ does not appear in the square of the Green's function.)

We now go back to the problem of the d-c conductivity. It was found that the important quantities for determining the d-c conductivity were the Fourier coefficients of the current-current correlation function in the restricted time region

$$\begin{aligned}
 F(t_1, t_2) &= \frac{e^2}{m} \int d^3 x_1 d^3 x_2 (-i)^2 T < \psi^+(\underline{x}_1, t_1) \nabla_1 \\
 &\times \psi(\underline{x}_1, t_1) \cdot \psi^+(\underline{x}_2, t_2) \nabla_2 \psi(\underline{x}_2, t_2) > \\
 &= \frac{e^2}{m} \int d^3 x_1 d^3 x_2 \nabla_1 \cdot \nabla_2 G(1, 2; 1', 2') \Big|_{\substack{1'=1 \\ 2'=2}}^+
 \end{aligned} \tag{304}$$

In order to get the physical picture we shall first make a rough approximation. Let us assume that the two-particle Green's function is factorizable:

$$G(1, 2; 1', 2') = G(1, 1') G(2, 2') - G(1, 2') G(2, 1') \tag{305}$$

Equation (305) corresponds essentially to the Hartree-Fock approximation. The first term in (305) will give nothing when substituted into (304), since the two single-particle Green's functions will be uncorrelated and we have translational symmetry. For the second term

$$F(\nu_m) = < \frac{e^2}{m^2} \sum_{\underline{k}, \underline{k}'} \frac{1}{\tau} \sum_l \underline{k} \cdot \underline{k}' G(\underline{k}, \underline{k}'; \zeta_l) G(\underline{k}', \underline{k}; \zeta_l - \nu_m) > \tag{306}$$

Substitute

$$G(\underline{k}, \underline{k}', \zeta_l) = \int \frac{d\omega}{2\pi} \frac{A(\underline{k}, \underline{k}', \omega)}{\zeta_l - \omega} \tag{307}$$

and do the sum over ζ_l by the usual technique of transforming to an integral where the contour envelopes the real axis. (Here both f^+ and f^- will give enough convergence.) Then (306) reduces to

$$F(\nu_m) = \frac{ie^2}{m^2} \sum_{\underline{k}, \underline{k}'} \underline{k} \cdot \underline{k}' \int \frac{d\omega_1 d\omega_2}{(2\pi)^2} \langle A(\underline{k}, \underline{k}', \omega_1) A(\underline{k}, \underline{k}', \omega_2) \rangle \cdot \frac{f^-(\omega_1) - f^-(\omega_2)}{\omega_1 - \omega_2 - \nu_m} \quad (308)$$

By using

$$f^-(\omega_1) - f^-(\omega_2) = \left(1 - e^{\beta(\omega_1 - \omega_2)}\right) f^-(\omega_1) f^+(\omega_2) \quad (309)$$

and

$$F(\nu_m) = i \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{\varphi(\omega)}{\nu_m - \omega} \quad (257)$$

we get

$$\varphi(\omega) = \frac{e^2}{m^2} \sum_{\underline{k}, \underline{k}'} \underline{k} \cdot \underline{k}' \int \frac{d\omega_1 d\omega_2}{(2\pi)^2} \langle A(\underline{k}, \underline{k}', \omega_1) A(\underline{k}, \underline{k}', \omega_2) \rangle_{\text{imp}} \cdot (1 - e^{-\beta\omega}) f^-(\omega_1) f^+(\omega_2) 2\pi \delta(\omega + \omega_1 - \omega_2) \quad (310)$$

It was shown earlier that the d-c conductivity was given by

$$\sigma(0) = \lim_{\omega \rightarrow 0} \frac{\varphi(\omega)}{\omega} \frac{1}{\delta \text{ vol}}$$

In the limit $\omega \rightarrow 0$, $\omega_1 \rightarrow \omega_2$ due to the δ function and

$$(1 - e^{-\beta\omega}) \rightarrow \beta\omega$$

if we now calculate in the low-temperature limit,

$$\lim_{\beta \rightarrow \infty} \beta f^+(\omega) f^-(\omega) = \delta(\omega - \mu) \quad (311)$$

so that

$$\sigma(0) = \frac{e^2}{12 \pi m^2 \text{Vol}} < \sum_{\underline{k} \underline{k}'} \underline{k} \cdot \underline{k}' A(\underline{k}, \underline{k}', \mu) A(\underline{k}', \underline{k}, \mu) >_{\text{imp}} \quad (312)$$

Let us for the moment disregard the difference between the product average and the average product. Then

$$\begin{aligned} \sigma &= \frac{e^2}{12 \pi m^2 \text{Vol}} \sum_{\underline{k}} \left(\frac{\Gamma_{\underline{k}}(\mu)}{(\mu - \epsilon_{\underline{k}} - \Delta_{\underline{k}}(\mu))^2 + \Gamma_{\underline{k}}^2(\mu)/4} \right)^2 k^2 \\ &\approx \frac{e^2}{12 \pi m^2 \text{Vol}} \sum_{\underline{k}} \frac{4 \pi k^2}{\Gamma_{\underline{k}}(\mu)} \delta(\mu - \epsilon_{\underline{k}} - \Delta_{\underline{k}}(\mu)) \end{aligned} \quad (313)$$

To see the last step in (313) note that

$$\frac{s}{x^2 + (s^2/4)}$$

as a function of x has a maximum at $x = 0$ with height $1/s^2$ and width s . If one works out the proper factors one sees that

$$\left(\frac{s}{x^2 + s^2/4} \right)^2 \simeq \frac{4\pi}{s} \delta(x) + O(s) \quad (314)$$

For a noninteracting system, (313) reduces to

$$\sigma_0 = \frac{n e^2}{m \Gamma_{\underline{k}_F}} \quad (315)$$

Even for a noninteracting system the result (315) is not what one would expect. In the Boltzmann equation approach discussed in Sec. 1 there was an additional contribution to the scattering rate from "scattering in" events. This is missing in (315) and (313). However, a more careful analysis of (312) will show that this contribution is present. The expression for the d-c conductivity thus obtained will still not be the correct one, the fault now lying in the ad hoc approximation (305). Before entering a discussion of how to improve (305) we give a detailed calculation of (312). (4) We had

$$\sigma = \frac{e^2}{12 \pi m^2 \text{Vol}} < \sum_{\underline{k}, \underline{k}'} \underline{k} \cdot \underline{k}' A(\underline{k}, \underline{k}', \omega) A(\underline{k}', \underline{k}, \omega) >_{\text{imp}} \quad (312)$$

From (97)

$$-iA(\underline{k}, \underline{k}', \omega) = G(\underline{k}, \underline{k}', \omega + i\epsilon) - G(\underline{k}, \underline{k}', \omega - i\epsilon) \quad (316)$$

It is a consequence of time-reversal invariance that

$$G(\underline{k}, \underline{k}', t) = G(-\underline{k}', -\underline{k}, t) \quad (317)$$

To see this note that time-reversal invariance essentially tells us that we can construct† an operator τ :

$$\begin{aligned} \tau a_{\underline{k}}(t) \tau^{-1} &= a_{-\underline{k}}(-t) \\ \tau a_{\underline{k}}^+(t) \tau^{-1} &= a_{-\underline{k}}^+(-t) \end{aligned} \quad (318)$$

† We have suppressed spin indices. In configuration space τ is the complex conjugation operator, i. e., the prescription is to take the complex conjugate of

$$a_{\underline{k}}(t) = e^{iHt} (\text{Vol})^{-1/2} \int \psi(\underline{x}) e^{i\underline{k} \cdot \underline{x}} d^3x e^{-iHt}$$

If we include spin $\tau = \sigma_y \times$ complex conjugation operator, which then has the additional effect $\sigma \rightarrow -\sigma$.

which has the properties

$$\langle \tau \psi_1 | \tau \psi_2 \rangle = \langle \psi_1 | \psi_2 \rangle^* = \langle \psi_2 | \psi_1 \rangle \quad (319)$$

and which commutes with the Hamiltonian of the system. Using (318) and (319) we get

$$\begin{aligned} \langle a_{\underline{k}}(t) a_{\underline{k}'}^+(0) \rangle &= \langle a_{-\underline{k}}(-t) a_{-\underline{k}'}^+(0) \rangle^* \\ &= \langle a_{-\underline{k}'}(0) a_{-\underline{k}}^+(-t) \rangle \\ &= \langle a_{-\underline{k}'}(t) a_{-\underline{k}}^+(0) \rangle \end{aligned} \quad (320)$$

We have shown that (317) holds for $G_>$, similar arguments give the same relation for $G_<$ and (317) therefore holds in general.

By substituting (316) and (317) into (312) we see that what we have to calculate is of the form

$$\begin{aligned} &\langle \sum_{\underline{k}, \underline{k}'} \underline{k} \cdot \underline{k}' [G(\underline{k}, \underline{k}', \mu + i\epsilon) - G(\underline{k}, \underline{k}', \mu - i\epsilon)] \\ &\cdot [G(-\underline{k}, -\underline{k}', \mu + i\epsilon) - G(-\underline{k}, -\underline{k}', \mu - i\epsilon)] \rangle_{\text{imp}} \end{aligned} \quad (321)$$

Let us define

$$K(\underline{k}, \omega_1, \omega_2) = \sum_{\underline{k}'} \underline{k} \cdot \underline{k}' \langle G(\underline{k}, \underline{k}', \omega_1) G(-\underline{k}, -\underline{k}', \omega_2) \rangle_{\text{imp}} \quad (322)$$

and

$$K^{(++)}(\underline{k}) = K(\underline{k}, \mu + i\epsilon, \mu + i\epsilon) \quad (323)$$

$$K^{(+-)}(\underline{k}) = K(\underline{k}, \mu + i\epsilon, \mu - i\epsilon) \quad (324)$$

In this notation [using the reality of $A(\underline{k}, \underline{k}', \omega)$ occurring in (316)]

$$\sigma = - \frac{2e^2}{12 \pi m^2 \text{Vol}} \sum_{\underline{k}} [\text{Re } K^{(++)}(\underline{k}) - K^{(+-)}(\underline{k})] \quad (325)$$

The quantity $K(\underline{k})$ satisfies an integral equation (essentially the Bethe-Salpeter equation). We first write down the Bethe-Salpeter equation and then justify it in terms of the graphs that occur in the impurity averaging:

$$K(\underline{k}, \omega_1, \omega_2) = k^2 \bar{G}(\underline{k}, \omega_1) \bar{G}(-\underline{k}, \omega_2) + \sum_{\underline{k}'} \bar{G}(\underline{k}, \omega_1) \bar{G}(-\underline{k}, \omega_2) \cdot \int \frac{d^3 l}{(2\pi)^3} W(\underline{k} - \underline{l}, \omega_1, \omega_2) \underline{k} \cdot \underline{k}' < G(\underline{l}, \underline{k}, \omega_1) G(-\underline{l}, -\underline{k}', \omega_2) >_{\text{imp}} \quad (326)$$

Graphically the structure of (326), apart from the common factor $\underline{k} \cdot \underline{k}'$, is given in Fig. 17. The left-hand side of Fig. 17 is pictured as a black

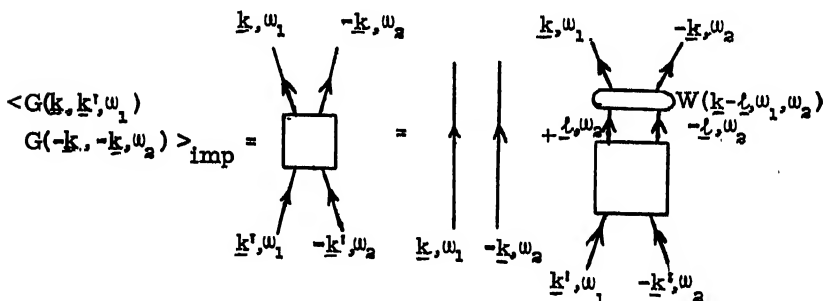


Fig. 17

box, containing an arbitrary number of elastic scattering processes of the electrons against the impurities. The graphs that occur fall into either of two classes. In the first type of graphs the two electrons do not scatter against the same impurity. This situation is indicated by the two straight lines in Fig. 17; in Eq. (326) the impurity average for the single-particle Green's function is indicated by a bar.

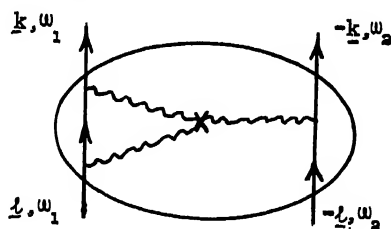


Fig. 18

A typical graph, contributing to the second term in Fig. 17, is given in Fig. 18, where the part that corresponds to W is encircled. Note that the structure of W is quite analogous to that of Σ in the case of the single-particle Green's function [see Eq. (20)]. W is the sum of all graphs in Fig. 18 of the type that are irreducible. By irreducible we here mean graphs that cannot be separated into two parts by cutting two particle lines. At the bottom of Fig. 18 we then have to attach all possible reducible and irreducible graphs, as indicated in Fig. 17. In the last term in (326) there occur three different directions in momentum space. Since

$$\begin{aligned} \cos(\underline{k}, \underline{k}') &= \cos(\underline{k}, \underline{l}) \cos(\underline{k}', \underline{l}) + \sin(\underline{k}, \underline{l}) \sin(\underline{k}', \underline{l}) \\ &\times \cos(\varphi_{\underline{k}} - \varphi_{\underline{k}'}), \end{aligned} \quad (327)$$

and since the last term in (327) vanishes when integrating over \underline{k}' we can put

$$\begin{aligned} K(\underline{k}, \omega_1, \omega_2) &= \overline{G}(\underline{k}, \omega_1) \overline{G}(-\underline{k}, \omega_2) \left[k^2 + \int_0^\infty l^2 dl \frac{k}{l} \right. \\ &\times w(k, l, \omega_1, \omega_2) K(l, \omega_1, \omega_2) \left. \right] \end{aligned} \quad (328)$$

where

$$w(k, l, \omega_1, \omega_2) = \int \frac{d\Omega_l}{(2\pi)^3} \cos(\underline{k}, \underline{l}) W(\underline{k} - \underline{l}, \omega_1, \omega_2) \quad (329)$$

The d-c conductivity was given by Eq. (325) and in order to say something about the conductivity we must solve Eq. (328).

Let us first evaluate K^+ . To lowest order in the impurity density

$$\begin{aligned} W(\underline{k} - \underline{l}, \mu + i\epsilon, \mu + i\epsilon) &= n_i \langle \underline{k} | \tilde{v}(\mu) | \underline{l}^+ \rangle \langle -\underline{k} | \tilde{v}(\mu) | -\underline{l}^- \rangle \\ &= n_i |t(\underline{k}, \underline{l}, \mu)|^2 \end{aligned} \quad (330)$$

The last step follows from time-reversal invariance. From (330)

$$w(\mathbf{k}, \ell, \mu + i\epsilon, \mu - i\epsilon) = n_i \int \frac{d\Omega}{(2\pi)^3} \cos(\hat{\mathbf{k}}, \hat{\ell}) |t(\underline{\mathbf{k}}, \underline{\ell}, \mu)|^2 \quad (331)$$

Equation (331) is very closely related to the expression we had for $\Gamma_{\mathbf{k}}(\mu)$:

$$\Gamma_{\mathbf{k}}(\mu) = 2\pi n_i \int \frac{d^3 p}{(2\pi)^3} |t(\underline{\mathbf{k}}, \underline{\mathbf{p}}, \mu)|^2 \delta(\mu - \epsilon_{\mathbf{p}} - \Delta_{\mathbf{p}}(\mu)) \quad (284)$$

Using

$$\delta(\mu - \epsilon_{\underline{\mathbf{p}}} - \Delta_{\underline{\mathbf{p}}}(\mu)) = \delta(\mu - \epsilon_{\underline{\mathbf{p}}}) = \sum_{\underline{\mathbf{p}}} \delta(\mu - \epsilon_{\underline{\mathbf{p}}}) \quad (332)$$

we get

$$\Gamma_{\mathbf{k}}(\mu) = \frac{2\pi n_i \sum_{\mathbf{k}_F} k_F^2}{v_F} \int \frac{d\Omega}{(2\pi)^3} |t(\underline{\mathbf{k}}, \underline{\mathbf{p}}_F, \mu)|^2 \quad (333)$$

which is proportional to (331) except for the cosine term. Define

$$\Gamma'_{\mathbf{k}}(\mu) = \frac{2\pi n_i \sum_{\mathbf{k}_F} k_F^2}{v_F} \int \frac{d\Omega}{(2\pi)^3} |t(\underline{\mathbf{k}}, \underline{\mathbf{p}}_F, \mu)|^2 \cos(\underline{\mathbf{k}}, \underline{\mathbf{p}}_F) \quad (334)$$

then

$$w(k, k_F, \mu + i\epsilon, \mu - i\epsilon) = \frac{1}{2\pi} \frac{v_F}{k_F^2 Z_{k_F}} \Gamma'_k(\mu) \quad (335)$$

From Eq. (326)

$$K^{+-}(k) = \frac{1}{Z_{k_F}^{-2} (\mu - E_k)^2 + \frac{\Gamma_k^2(\mu)}{4}} \times \left\{ k^2 + \int_0^\infty d\ell \, k\ell \, w^{+-}(k, \ell) K^{+-}(\ell) \right\} \quad (336)$$

In order to solve (336) it is convenient to introduce the quantity

$$\begin{aligned} M &\equiv \int_0^\infty d\ell \, \ell \, w^{+-}(k_F, \ell) K^{+-}(\ell) \\ &= \int_0^\infty d\ell \, \frac{\ell^3 w^{+-}(k_F, \ell)}{\alpha_\ell^2 + \Gamma_\ell^2/4} + \int_0^\infty d\ell \, \left(\frac{\ell^2 w_{+-}(k_F, \ell)}{\alpha_\ell^2 + \Gamma_\ell^2/4} \right. \\ &\quad \left. \times \int_0^\infty d\ell' \, \ell' w^{+-}(\ell\ell') K^{+-}(\ell') \right) \end{aligned} \quad (337)$$

where

$$\alpha_k^2 = Z_k^{-2} (\mu - E_k)^2 \quad (338)$$

The function

$$\frac{w_{+-}(k_F, \ell)}{\alpha_\ell^2 + (\Gamma_\ell^2/4)} \quad (339)$$

here plays the role of a δ function. The normalization is

$$\int_0^{\infty} d\ell \frac{w_{+-}(k_F, \ell)}{\alpha_{\ell}^2 + \Gamma_{\ell}^2 / 4} = w_{+-}(k_F, k_F) \frac{Z_{k_F}}{v_F} \frac{2\pi}{\Gamma_{k_F}} \quad (340)$$

so that

$$\frac{w_{+-}(k_F, \ell)}{\alpha_{\ell}^2 + \Gamma_{\ell}^2 / 4} = \frac{\Gamma_{k_F}'}{k_F^2 \Gamma_{k_F}} \delta(\ell - k_F) + O\left(\frac{\Gamma}{\mu}\right) \quad (341)$$

Solving for M

$$M = \frac{k_F \Gamma_{k_F}'}{\Gamma_{k_F}} + \frac{\Gamma_{k_F}' M}{\Gamma_{k_F}} \quad (342)$$

or

$$M = \frac{k_F \Gamma_{k_F}'}{\Gamma_{k_F} - \Gamma_{k_F}'} \quad (343)$$

We therefore have

$$K^{+-}(k) = \frac{1}{\alpha_k^2 + \Gamma_k^2 / 4} \left[k^2 + \frac{k k_F \Gamma_{k_F}'}{\Gamma_{k_F} - \Gamma_{k_F}'} \right] \quad (344)$$

Going back to (325) it is seen that (344) gives a contribution to the conductivity that diverges for large k . It can be shown⁽³⁾ that the contribution from K^{++} exactly cancels this divergence, whereas around $k = k_F$, $K^{++}(k)$ is of the order Γ/μ and can be neglected. We therefore only have to consider

$K^{+-}(k)$ for $k \sim k_F$ where

$$K^{+-}(k) \simeq \frac{2\pi}{v_F} \frac{Z_{k_F}}{\Gamma_{k_F}} \frac{k_F^2 \Gamma_{k_F}}{\Gamma_{k_F} - \Gamma_{k_F}} \delta(k - k_F) \quad (345)$$

and we finally get

$$\sigma = \frac{ne^2}{m} \left(\frac{k_F Z_{k_F}}{v_F m} \right) \frac{1}{\Gamma_{k_F} - \Gamma_{k_F}} \quad (346)$$

Equation (346) is still not quite the correct answer. We shall see that in order to get the correct expression we shall have to include vertex corrections. Before we do this a few remarks on renormalization are in order. We had for the single-particle Green's function

$$G(\underline{k}, \mu + i\epsilon) = \frac{1}{(\mu - \epsilon_{\underline{k}}) + i\Gamma_{\underline{k}}/2} \quad (347)$$

In terms of the quasi-particle energies $E_{\underline{k}}$ we get

$$G(\underline{k}, \mu + i\epsilon) = \frac{Z_{\underline{k}}}{(\mu - E_{\underline{k}}) + i\Gamma_{\underline{k}}/2} \quad (348)$$

It is seen that $Z_{\underline{k}}$ is the many-body analogue of the wave function renormalization constant. Similarly the S matrix

$$S_{\underline{k}\underline{k}'}(\mu) = \delta_{\underline{k}\underline{k}'} - 2\pi i \delta(\epsilon_{\underline{k}'} - \mu) t(\underline{k}, \underline{k}', \mu) \quad (349)$$

becomes in terms of the quasi-particle energies

$$S_{\underline{k}\underline{k}'}(\mu) = \delta_{\underline{k}\underline{k}'} - 2\pi i \delta(E_{\underline{k}'} - \mu) \bar{T}(\underline{k}, \underline{k}', \mu) \quad (350)$$

where $\bar{t}(\underline{k}, \underline{k}', \mu)$ is the renormalized t matrix

$$\bar{T}(\underline{k}, \underline{k}', \mu) = Z_{\underline{k}} t(\underline{k}, \underline{k}', \mu) \quad (351)$$

We can also express

$$\bar{\Gamma}_{\mathbf{k}_F} - \bar{\Gamma}'_{\mathbf{k}_F} = 2\pi n_i Z_{\mathbf{k}_F}^2 \int \frac{d^3 p}{(2\pi)^3} |t(\mathbf{p}_F, \theta)|^2 (1 - \cos \theta) \delta(\mu - E_p) \quad (352)$$

in terms of the renormalized quantities

$$\bar{\Gamma}_{\mathbf{k}_F} - \bar{\Gamma}'_{\mathbf{k}_F} = 2\pi n_i \int \frac{d^3 p}{(2\pi)^3} |\bar{T}(\mathbf{p}_F, \theta)|^2 (1 - \cos \theta) \delta(\mu - E_p) \quad (353)$$

The transport scattering time is defined by

$$\frac{1}{\tau_{tr}} = n_i v_F \int \frac{d\sigma}{d\Omega} (1 - \cos \theta) d\Omega \quad (354)$$

where $(d\sigma/d\Omega)$ is the differential cross section for impurity scattering. From (353)

$$\tau_{tr} = \frac{1}{\bar{\Gamma}_{k_F} - \bar{\Gamma}'_{k_F}} \quad (355)$$

i. e., the correct transport scattering time is also to be expressed in terms of the renormalized quantities. Defining the effective mass m^* by

$$\frac{k_F}{m^*} = v_F \quad (356)$$

and substituting (355) into (346)

$$\sigma = \frac{n e^2 \tau_{tr}}{m^*} \left(\frac{m^* Z_{k_F}}{m} \right)^2 \quad (357)$$

For the noninteracting case we have

$$\sigma = \frac{n e^2 \tau_{tr}}{m} \quad (358)$$

The Landau theory of Sec. 1 gave the result

$$\frac{n e^2 \tau_{tr}}{m^*} \quad (359)$$

In (357) we seem to have an additional multiplicative factor. In what follows we shall show that this factor is due to the approximate factorization (305) and that it does not appear when the vertex describing the interaction of the quasi-particles and the external field is properly treated.

The Reduced Graph Expansion

The approximate calculation described above was based on the ad hoc factorization (305) of the two-particle Green's function. We now show that it is possible to avoid this approximation and to generate instead a formally exact expansion. Fortunately, only the first term of this expansion will survive at $T = 0$ and we shall show that this term reduces to the expression (359). To begin, let us briefly review the general procedure for calculating σ . We have first to calculate the current correlation function

$$F(t, t') = T \langle \bar{j}(t) \cdot j(t') \rangle \quad (360)$$

and compute the Fourier series coefficients $F(\nu_m)$ for t, t' in the restricted time region. Then we have to compute $\varphi(w)$ the discontinuity of $F(\nu_m = z)$ across the real axis. [By $F(\nu_m = z)$ we mean that continuation to the complex z plane which has as its only singularities a branch cut along the real axis.] Finally we have the rule

$$\sigma = \frac{1}{6(Vol)} \lim_{w \rightarrow 0} \frac{\varphi(w)}{w} \quad (253)$$

We can carry out the above in a formally exact way by first making a perturbation expansion of $F(\nu_m)$ and then calculating the discontinuity for each term in the perturbation series. The formal technique we shall use was first developed by Langer.⁽¹⁴⁾ It is based on methods developed by Landau⁽¹⁵⁾ and Cutkosky⁽¹⁶⁾ for relativistic quantum field theory.

It is a straightforward application of the perturbation theory discussed in Sec. 2 to generate a graphical expansion for $F(\nu_m)$. The rules thus derived may be summarized as follows. Draw all skeleton graphs (i. e.,

graphs in which no electron line has any self-energy insertions) that are linked and have two external vertices and with each graph associate the following factors:

- (1) $iv(q)$ for each explicit internal interaction line carrying momentum \underline{q}

$$(2) G_{\underline{k}\underline{k}'}(\zeta_{\ell}) = \int_{-\infty}^{\infty} \frac{A_{\underline{k}\underline{k}'}(\omega)}{\zeta_{\ell} - \omega} \frac{d\omega}{2\pi}$$

for each electron line having the momentum and quasi-energy labels indicated

- (3) (-1) for each closed electron loop

- (4) $-(e^2 \underline{k} \cdot \underline{k}' / m^2)$ for the external vertices at which electrons with momenta \underline{k} and \underline{k}' are scattered with no momentum transfer

Momentum and quasi energy are to be conserved at each vertex. The external vertices have a quasi-energy transfer ν_m . Finally, sum over all free quasi-energy variables $(1/\tau)\Sigma_{\ell}$ and all free momenta $\int d^3 k / (2\pi)^3$.

To calculate the contribution to $\varphi^{(n)}$ from a particular graph one can proceed in three steps.

- (1) Hold the ω 's fixed and isolate those combinations of factors $1/(\zeta_{\ell} - \omega)$ that will give rise to poles in the ν_m plane
- (2) Calculate the residues at these poles.
- (3) Perform the ω integrations generating branch cuts (more than one in general) along the real axis in the ν_m plane, and calculate the total discontinuity across the real axis.

To illustrate this procedure we consider the graph of Fig. 19. According to our rules the contribution of this graph to $F(\nu)$ is

$$F_G(\nu) = 4(i)^2 (-1)^2 \frac{e^2}{m^2} \\ \times \sum_{\substack{\underline{k}\underline{k}'\underline{q} \\ \underline{p}_1 \underline{p}_2 \underline{p}_3}} (-\underline{k} \cdot \underline{k}') |v(q)|^2 \int \frac{d\omega_1 \dots d\omega_6}{(2\pi)^6}$$

$$\begin{aligned}
& \times \frac{1}{r^3} \sum_{\zeta_1 \dots \zeta_6} \frac{A_{\underline{k}, \underline{p}_1, \underline{q}}^{(\omega_1)} A_{\underline{p}, \underline{k}}^{(\omega_2)} A_{\underline{p}_2, \underline{p}_3}^{(\omega_3)}}{(\zeta_1 - \omega_1) (\zeta_2 - \omega_2) (\zeta_3 - \omega_3)} \\
& \times \frac{A_{\underline{p}_3, \underline{q}, \underline{p}_2, \underline{q}}^{(\omega_4)} A_{\underline{k}, \underline{p} + \underline{q}}^{(\omega_5)} A_{\underline{p}_1, \underline{k}}^{(\omega_6)}}{(\zeta_4 - \omega_4) (\zeta_5 - \omega_5) (\zeta_6 - \omega_6)} \\
& \times \delta_{\zeta_1 + \nu, \zeta_2} \delta_{\zeta_2 + \zeta_3, \zeta_4 + \zeta_5} \delta_{\zeta_4 + \zeta_6, \zeta_3 + \zeta_1} \quad (361)
\end{aligned}$$

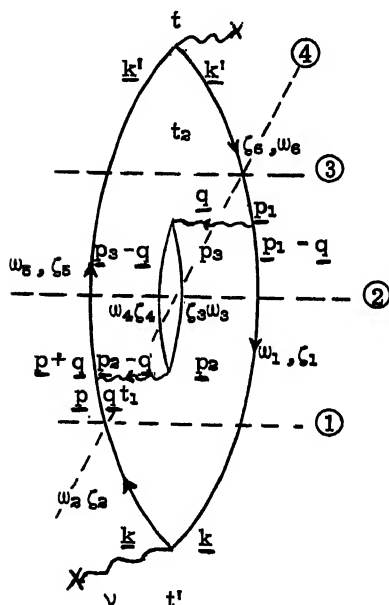


Fig. 19

The first step in our prescription for finding the contribution to $\varphi(\omega)$ is best done in the time representation, i. e., we use the representations

$$\delta_{\zeta_2+\zeta_3, \zeta_4+\zeta_5} = \frac{1}{\tau} \int_0^\tau e^{-it(\zeta_2+\zeta_3-\zeta_4-\zeta_5)} dt \quad (362)$$

and

$$\begin{aligned} & \frac{1}{\tau} \sum_{\zeta} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{A_{kk'}(\omega)}{\zeta - \omega} e^{-i\zeta t} \\ &= \begin{cases} t > 0 & \frac{1}{i} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} A_{kk'}(\omega) f^+(\omega) e^{-i\omega t} \\ t < 0 & -\frac{1}{i} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} A_{kk'}(\omega) f^-(\omega) e^{-i\omega t} \end{cases} \quad (363) \end{aligned}$$

for each of the internal interaction lines and each of the electron lines, respectively. Furthermore for one external vertex we use the representation

$$\delta_{\zeta_1 - \zeta_2 + \nu} = \frac{1}{\tau} \int_0^\tau e^{-it(\zeta_1 - \zeta_2 + \nu)} dt \quad (364)$$

When this is done, the contribution of the topological graph Fig. 19 breaks up naturally into a sum of contributions from the various possible orderings in time of the interaction lines. The time ordering illustrated, for example, has the integrations

$$\begin{aligned}
& e^{-i(\omega_5 - \omega_6 - \nu)t} \int_0^t dt_2 e^{-i(\omega_4 + \omega_6 - \omega_3 - \omega_1)t_2} \\
& \times \int_0^{t_2} dt_1 e^{-i(\omega_2 + \omega_3 - \omega_4 - \omega_5)t_1} \int_0^{t_1} dt' e^{-i(\nu + \omega_1 - \omega_2)t'}
\end{aligned}
\tag{365}$$

In doing these integrals, the following denominators involving ν occur:

$$\begin{aligned}
(1) \quad & \frac{1}{\nu + \omega_1 - \omega_2} \\
(2) \quad & \frac{1}{\nu + \omega_1 + \omega_3 - \omega_5 - \omega_6} \\
(3) \quad & \frac{1}{\nu + \omega_6 - \omega_5}
\end{aligned}$$

These poles correspond to nothing more than the simultaneous vanishing of certain "energy denominators," taking into account quasi-energy conservation. For example pole (1) is a solution of

$$\omega_2 = \zeta_2; \quad \omega_1 = \zeta_1; \quad \nu = \zeta_2 - \zeta_1
\tag{366}$$

Pole (2) is a solution of

$$\omega_{1,3,5,6} = \zeta_{1,3,5,6}; \quad \zeta_5 + \zeta_4 - \zeta_3 - \zeta_1 = \nu
\tag{367}$$

Similarly pole (3) is a solution of

$$\omega_{5,6} = \zeta_{5,6}; \quad \zeta_5 - \zeta_6 = \nu \quad (368)$$

In other time orderings, one other denominator containing ν arises. This corresponds to the simultaneous solution of

$$\omega_{2,4,3,6} = \zeta_{2,4,3,6}; \quad \zeta_2 + \zeta_3 - \zeta_4 - \zeta_6 = \nu \quad (369)$$

The position of this pole [call it pole (4)] is thus given by

$$(4) \quad \nu + \omega_6 + \omega_4 - \omega_3 - \omega_2 = 0$$

The singularities in the ν plane thus come from the simultaneous satisfaction of the conditions $\zeta_l = \omega_l$ for all lines that in some time ordering can be cut by a horizontal line, provided that the lines so cut have quasi energies related by the conservation at the vertices to ν . (The cut indicated in Fig. 20, which is a particular time-ordered constituent of Fig. 19, gives no singularity in the ν plane.) Actually the prescription given above leads to some spurious poles. A slightly more

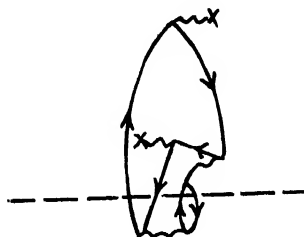


Fig. 20

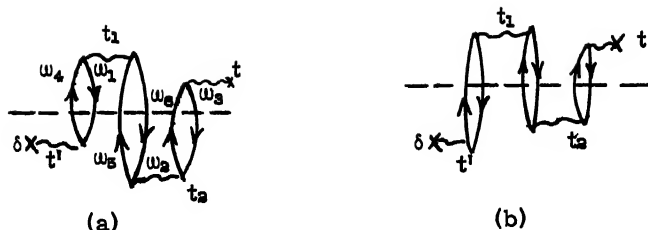


Fig. 21

precise statement is needed. To see the difficulty consider the graphs of Fig. 21. The time ordering illustrated in Fig. 21(a) leads to a denominator

$$\frac{1}{\nu + \omega_1 + \omega_2 + \omega_3 - \omega_4 - \omega_5 - \omega_6} \quad (370)$$

This pole arises because the upper limit of the integration over t_2 is t' in the graph of Fig. 21(a). However in the graph of Fig. 21(b) in which $t_2 > t'$ the integrand is identical to that of Fig. 21(a) because the hole or particle character of the lines has nowhere been changed. Therefore the restriction $t_2 \leq t'$ simply does not exist and the pole (370) disappears. In this way we find that the only poles of the graph of which Fig. 21(a) and (b) are time-ordered constituents are given by $\nu = \omega_4 - \omega_1$, $\nu = \omega_5 - \omega_2$, $\nu = \omega_6 - \omega_3$. The

general rule, then, for finding which combinations of denominators lead to poles in the ν plane is this. The poles come from the simultaneous satisfaction of $\zeta_l = \omega_l$ for all lines that in some time ordering can be cut by a

horizontal line that divides the graph into two portions each of which is connected to one vertex. The cuts that give rise to the four poles coming from Fig. 19 are indicated on the figure. We have thus been able to perform part (1) of the three-step prescription for finding the contribution to $\varphi(\omega)$ from a particular graph.

We have now to perform step (2), i. e., to find what the residue is at the poles. To do this we need the identity

$$\frac{1}{\tau^{2p-1}} \sum_{\zeta_1 \dots \zeta_{2p}} \frac{1}{\zeta_1 - \omega_1} \frac{1}{\zeta_2 - \omega_2} \dots \frac{1}{\zeta_{2p} - \omega_{2p}} \\ \times \delta_{\zeta_1 + \zeta_2 \dots \zeta_p + \nu, \zeta_{p+1} + \dots \zeta_{2p}}$$

$$= \frac{(-i)^{2p-1} (-1)^p f^-(w_1) f^-(w_2) \dots f^-(w_p) f^+(w_{p+1}) \dots f^+(w_{2p}) \left[1 - e^{\beta(w_1 + w_2 \dots - w_{2p})} \right]}{v + w_1 + w_2 + \dots w_p - w_{p+1} \dots - w_{2p}}$$

(371)

This is simply proved using the contour integral technique. Defining

$$\begin{aligned} \zeta_1 &= v_1 + \mu + \frac{\pi}{\tau} & v_1 &= \frac{2\pi}{\tau} \times (\text{even integer}) \\ \zeta_2 &= v_2 - v_1 + \mu + \frac{\pi}{\tau} \\ &\vdots \\ \zeta_p &= v_p - v_{p-1} + \mu + \frac{\pi}{\tau} \\ \zeta_{p+1} &= v_{p+1} + v_p + \mu + \frac{\pi}{\tau} \\ \zeta_{p+2} &= v_{p+2} - v_{p+1} + \mu + \frac{\pi}{\tau} \\ &\vdots \\ \zeta_{2p} &= v - v_{2p-1} + \mu + \frac{\pi}{\tau} \end{aligned} \tag{372}$$

one automatically satisfies the δ condition in (371). Doing the sums over $v_1 \dots v_{2p-1}$ successively using

$$\frac{1}{\tau} \sum_v \dots = \frac{(-i)}{2\pi i} \oint_{\Gamma_0} dv \frac{1}{e^{i\tau v} - 1} \dots \quad (373)$$

and using the fact that $e^{i\tau v_i} = 1$ leads directly to the right-hand side of (371). Using the theorem (371) we are able to write down at once the residues at the poles which we have learnt to enumerate in step (1) of our procedure. One has simply to use (371) to combine those energy denominators that give rise to poles. The remaining part of the integrand is to be evaluated at the position of the pole. Consider again the graph of Fig. 19. The quasi-energy conservation conditions are automatically satisfied by a new choice of variables

$$\zeta_1 = v_1 + \mu + \frac{\pi}{\tau}$$

$$\zeta_2 = v + v_1 + \mu + \frac{\pi}{\tau}$$

$$\zeta_3 = v_2 - v_1 + \mu + \frac{\pi}{\tau}$$

$$\zeta_4 = v_3 + v_2 + \mu + \frac{\pi}{\tau}$$

$$\zeta_5 = v - v_3 + \mu + \frac{\pi}{\tau}$$

$$\zeta_6 = -v_3 + \mu + \frac{\pi}{\tau} \quad (374)$$

The "energy denominators" for the whole graph are then

$$\begin{aligned}
& \frac{1}{\tau^3} \sum_{\substack{\nu_1 \nu_2 \\ \nu_3}} \frac{1}{\nu_1 - \omega_1 + \mu + \frac{\pi}{\tau}} \cdot \frac{1}{\nu + \nu_1 - \omega_2 + \mu + \frac{\pi}{\tau}} \cdot \frac{1}{\nu_2 - \nu_1 - \omega_3 + \mu + \frac{\pi}{\tau}} \\
& \cdot \frac{1}{\nu_3 + \nu_2 - \omega_4 + \mu + \frac{\pi}{\tau}} \cdot \frac{1}{\nu - \nu_3 - \omega_5 + \mu + \frac{\pi}{\tau}} \cdot \frac{1}{-\nu_3 - \omega_6 + \mu + \frac{\pi}{\tau}} \quad (375)
\end{aligned}$$

For pole (1) we have to combine the denominators containing ω_1 and ω_2 . Using (371) the residue at this pole is

$$\begin{aligned}
& (-i) (-1) f^-(\omega_1) f^+(\omega_2) [1 - e^{\beta(\omega_1 - \omega_2)}] \\
& \times \left[\frac{1}{\tau^2} \sum_{\nu_2 \nu_3} \frac{1}{\nu_2 - \nu_1 - \omega_3 + \mu + \frac{\pi}{\tau}} \dots \right] \\
& \nu_1 = \omega_1 - \mu - \frac{\pi}{\tau} \\
& \nu = \omega_2 - \omega_1 \quad (376)
\end{aligned}$$

where the dots in the square bracket indicate the three remaining denominators. The sums over ν_2 and ν_3 are to be done before the substitutions for ν_1 and ν are made in order to avoid the ambiguity we have repeatedly discussed.

We have thus indicated how to carry out step (2) of the procedure for calculating the contribution in perturbation theory to $\varphi(\omega)$. The third step is now quite trivial. Carrying out the ω integrations one generates a branch cut along the axis $\text{Im } \nu = 0$ for each pole. The discontinuity from a term of the form $1/(\nu - a)$ (where a is some linear combination of ω_i 's) at the point $\text{Re } \nu = \omega$ is simply $-2\pi i \delta(\omega - a)$. The only slight complication that arises is that in general one has overlapping branch cuts. What to do in this case was, however, carefully discussed in the derivation of the

optical theorem given before, Eq. (282). One has simply to order the branch cuts in some definite way, displace them slightly from each other, and then cross them one by one.

One can now make an expansion of the discontinuity function by putting together all contributions that arise from the vanishing of a particular combination of energy denominators. This is the so-called reduced graph expansion. One represents the terms by drawing graphs in which only the lines corresponding to vanishing energy denominators are exhibited, the rest being fused together at the vertices. For example for $\varphi(\omega)$ the two lowest-order reduced graphs are drawn in Fig. 22.

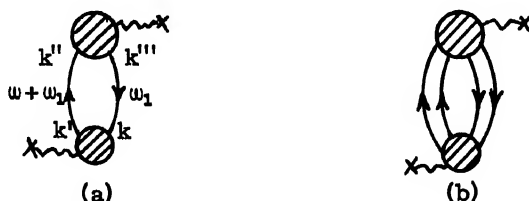


Fig. 22

Note that the particular example of Fig. 18 contributes to both these reduced graphs but to no others. Poles (1) and (3) of this example have intermediate states with one electron and one hole. These contribute to the reduced graph (a); poles (2) and (4) contribute to the reduced graph (b). Making use of the rules we have developed, one finds for the contribution of the reduced graph of Fig. 22(a):

$$\begin{aligned} \varphi^{(1)}(\omega) = & \left\langle \sum_{\substack{\mathbf{k}, \mathbf{k}' \\ \mathbf{k}, \mathbf{k}''}} \int_{-\infty}^{\infty} \frac{d\omega_1}{2\pi} A_{\mathbf{k}, \mathbf{k}'}(\omega + \omega_1 - i\eta, \omega_1) \right. \\ & \cdot A_{\mathbf{k}, \mathbf{k}'}(\omega_1, \omega + \omega_1 + i\eta) A_{\mathbf{k}, \mathbf{k}'}(\omega + \omega_1) A_{\mathbf{k}, \mathbf{k}'}(\omega_1) f^-(\omega_1) f^+(\omega + \omega_1) \\ & \left. \times [1 - e^{-\beta\omega}] \right\rangle_{\text{imp}} \end{aligned} \quad (377)$$

where the Λ 's are the vertex functions indicated by the shaded circles in Fig. 22(a). The factors $i\eta$ (η is an infinitesimal real positive quantity) are inserted in order to insure that in crossing the branch cut associated with a one-electron hole pair excitation we remain below (above) any branch cuts belonging to the upper (lower) vertex function. The contribution of the reduced graph of next order, i.e., Fig. 22(b) may be written schematically

$$\begin{aligned} \varphi^\omega(\omega) = & \left\langle \sum_{\text{mom}} \int \frac{d\omega_1 d\omega_2 d\omega_3}{(2\pi)^3} \Lambda^2(\omega - i\eta) \Lambda \Lambda \Lambda \Lambda \right. \\ & \times \Lambda^2(\omega + i\eta) f^-(\omega_1) f^-(\omega_2) f^+(\omega_3) f^+(\omega + \omega_1 + \omega_2 - \omega_3) \\ & \left. \times \left[1 - e^{-\beta\omega} \right] \right\rangle_I \end{aligned} \quad (378)$$

It is easy to see that the expansion we have generated leads to an expansion in even powers of the temperature for the d-c conductivity. We are interested in the limit of $\varphi(\omega)/\omega$ as $\omega \rightarrow 0$. This limit may be taken at once by replacing ω by zero everywhere but in the factor $[1 - e^{-\beta\omega}]$ and here using

$$\lim_{\omega \rightarrow 0} \frac{[1 - e^{-\beta\omega}]}{\omega} = \beta \quad (379)$$

For low temperatures the leading temperature dependence of (378) is obtained by considering the integral

$$\begin{aligned} & \beta \int \frac{d\omega_1}{2\pi} \frac{d\omega_2}{2\pi} \frac{d\omega_3}{2\pi} \\ & \times f^-(\omega_1) f^-(\omega_2) f^+(\omega_3) f^+(\omega_1 + \omega_2 - \omega_3) \propto \beta^{-2} \end{aligned} \quad (380)$$

Similarly the reduced graph having p electron and p hole lines is proportional to $\beta^{-(2p-2)}$. Thus only the reduced graph 22(a) survives as $T \rightarrow 0$, and from Eq. (377) we see that its contribution is

$$\sigma = \frac{1}{12\pi(\text{Vol})} \left\langle \sum_{\underline{k}\underline{k}', \underline{k}''} \underline{\Lambda}_{\underline{k}''}(\mu - i\eta) \cdot \underline{\Lambda}_{\underline{k}\underline{k}'}(\mu + i\eta) A_{\underline{k}\underline{k}'}(\mu) A_{\underline{k}''}(\mu) \right\rangle_I \quad (381)$$

Equation (381) can be further simplified. We are only interested in calculating σ to the order $(n_i)^{-1}$ where n_i is the density of impurities. Now in the absence of impurities the spectral weight function A contains a δ -function part at the Fermi surface. Therefore, in order to get a finite result at all, we must take into account the finite lifetime of the excitations. Once this is done, however, the additional impurity dependence will be of higher order. We may therefore compute the vertex functions to lowest order in the impurity density. Then (381) reduces to

$$\sigma = \frac{1}{12\pi(\text{Vol})} \sum_{\underline{k}\underline{k}'} \underline{\Lambda}_{\underline{k}}(\mu - i\eta) \cdot \underline{\Lambda}_{\underline{k}'}(\mu + i\eta) \cdot \left\langle A_{\underline{k}\underline{k}'}(\mu) A_{\underline{k}'}(\mu) \right\rangle_{\text{imp}} \quad (382)$$

Note that the only difference between (382) and the expression we had earlier in a Hartree-Fock approximation is in the appearance of the vertex function $\underline{\Lambda}$ instead of $(e/m)\underline{k}$. The vertex function is the sum of all graphs indicated in Fig. 23, and we have now to express it in terms of the parameters of the Landau theory.

Vertex Function and Ward Identity

We now proceed to calculate the vertex functions that occur in (382). We are interested in the limit of small momentum transfer, small energy transfer, and small impurity width. We shall see in a moment that one must be very careful in taking these limits and that the order in which they are taken is important. Before doing this we need a closed expression for the vertex function $\underline{\Lambda}(\underline{k}, \zeta_{\underline{k}}; \underline{q}, \nu_{\underline{n}})$ corresponding to momentum transfer \underline{q} and quasi-energy transfer $\nu_{\underline{n}}$. We see from Fig. 22(a) that $\underline{\Lambda}(\underline{k}, \zeta_{\underline{k}}; \underline{q}, \nu_{\underline{n}})$ is closely related to the quantity

$$T \left\langle a_{\underline{k}-\underline{q}}(t) \downarrow(\underline{q}, t') a_{\underline{k}}^{\dagger}(0) \right\rangle \quad (383)$$

except for the electron propagators at the ends, which we have to remove. The current operator is given by

$$\underline{j}(\underline{x}, t) = \frac{e}{2mi} \left[\psi^\dagger(\underline{x}, t) \underline{\nabla} \psi(\underline{x}, t) - \left(\underline{\nabla} \psi^\dagger(\underline{x}, t) \psi(\underline{x}, t) \right) \right] \quad (384)$$

or, in momentum space

$$\underline{j}(\underline{q}, t) = \frac{e}{m} \sum_{\underline{k}} \left(\underline{k} + \frac{\underline{q}}{2} \right) a_{\underline{k}}^+ a_{\underline{k}+\underline{q}} \quad (385)$$

The desired closed expression is, therefore,

$$\begin{aligned} & \sum_{\underline{k}'} \int_0^T dt \int_0^T dt' e^{i(\zeta_{\underline{k}'} - \nu_m)(t-t')} e^{i\zeta_{\underline{k}'} t'} (\underline{k}' + \underline{q}/2) \\ & \times T \left\langle a_{\underline{k}-\underline{q}}(t) a_{\underline{k}'}^+(t') a_{\underline{k}'+\underline{q}}(t') a_{\underline{k}}^+(0) \right\rangle \\ & = - G_{\underline{k}-\underline{q}}(\zeta_{\underline{k}} - \nu_m) \underline{\Lambda}(\underline{k}, \zeta_{\underline{k}}, \underline{q}, \nu_m) G_{\underline{k}}(\zeta_{\underline{k}}) \end{aligned} \quad (386)$$

Note that the vertex function represents the change in the single-particle Green's function due to an external disturbance.

When taking the desired limits it is important to note that a d-c measurement means that the frequency is small compared to the inverse of all the characteristic times of the system. In particular we must have

$$\nu \ll \Gamma \quad (387)$$

where $1/\Gamma$ is the relaxation time. This means that the desired order of the limits is

$$\lim_{\Gamma \rightarrow 0} \lim_{\nu \rightarrow 0} \lim_{q \rightarrow 0} \underline{\Lambda}^I(\underline{k}, \zeta_{\underline{k}}; \underline{q}, \nu) \quad (388)$$

i. e., we cannot ignore the impurities altogether. We will show, however, that

$$\begin{aligned} \lim_{\Gamma \rightarrow 0} \lim_{\nu \rightarrow 0} \lim_{q \rightarrow 0} \Lambda^I(k, \zeta_\ell) &= \lim_{q \rightarrow 0} \lim_{\nu \rightarrow 0} \lim_{\Gamma \rightarrow 0} \underline{\Lambda}^I(k, \zeta_\ell) \\ &= \lim_{q \rightarrow 0} \lim_{\nu \rightarrow 0} \underline{\Lambda}(k, \zeta_\ell) \end{aligned} \quad (389)$$

and thus after reversing the order of the limits we may ignore the impurities. Let us illustrate the preceding remarks by an example. Consider the skeleton graph (Fig. 23).

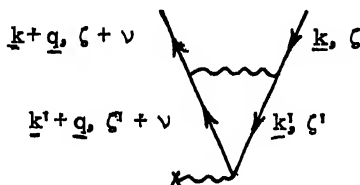


Fig. 23

The contribution to the vertex function from this graph is

$$\begin{aligned} \frac{1}{\tau} \sum_{\underline{k}', \zeta_\ell'} \int \frac{d\omega_1 d\omega_2}{(2\pi)^2} \frac{A_{\underline{k}'}(\omega_1) A_{\underline{k}'+\underline{q}}(\omega_2)}{(\zeta_\ell' - \omega_1)(\zeta_\ell' + \nu - \omega_2)} \\ \times \left(\underline{k}' + \frac{\underline{q}}{2} \right) v(\underline{k} - \underline{k}') \end{aligned} \quad (390)$$

Using the by now familiar contour integral techniques we get

$$\begin{aligned} - \sum_{\underline{k}'} v(\underline{k} - \underline{k}') \int \frac{d\omega_1 d\omega_2}{(2\pi)^2} A_{\underline{k}'}(\omega_1) A_{\underline{k}'-\underline{q}}(\omega_2) \\ \times \left(\underline{k}' + \frac{\underline{q}}{2} \right) \frac{f^-(\omega_1) - f^-(\omega_2)}{\nu + \omega_1 - \omega_2} \end{aligned} \quad (391)$$

For a pure system at $T = 0$ we have for $k \approx k_F$

$$A_{\underline{k}}(\omega) = 2\pi Z_{\underline{k}} \delta(\omega - E_{\underline{k}}) + \tilde{A}_{\underline{k}}(\omega) \quad (392)$$

Let us only consider the contribution from the peaked part

$$\Lambda_G = \sum_{\underline{k}'} Z_{\underline{k}'}^2 v(\underline{k} - \underline{k}') \frac{f^-(E_{\underline{k}'}) - f^-(E_{\underline{k}'+\underline{q}})}{v + E_{\underline{k}'} - E_{\underline{k}'+\underline{q}}} \left(\underline{k}' + \frac{\underline{q}}{2} \right) \quad (393)$$

For small \underline{q} Eq. (393) reduces to

$$-\sum_{\underline{k}'} \frac{Z_{\underline{k}'}^2 v(\underline{k} - \underline{k}') (E_{\underline{k}'+\underline{q}} - E_{\underline{k}'}) \delta(\mu - E_{\underline{k}'})}{v + E_{\underline{k}'} - E_{\underline{k}'+\underline{q}}} \quad (394)$$

and we see that

$$\lim_{q \rightarrow 0} \lim_{v \rightarrow 0} \lim_{\Gamma \rightarrow 0} \Lambda_G^I = -Z_{k_F}^2 \sum_{\underline{k}'} v(\underline{k} - \underline{k}') \delta(\mu - E_{\underline{k}'}) \underline{k}' \quad (395)$$

but

$$\lim_{v \rightarrow 0} \lim_{q \rightarrow 0} \lim_{\Gamma \rightarrow 0} \Lambda_G^I = 0 \quad (396)$$

This ambiguity is quite a general feature because any graph (see Fig. 24) that may be broken into two pieces by cutting one particle and one hole line



Fig. 24

will have a discontinuous behavior. The reason for this is that the spectral weight function in the absence of the impurities is very sharply peaked over the domain where the energy denominators vanish.

In order to satisfy the restriction $v \ll \Gamma$ it is necessary to round off the peak, i.e., to replace

$$Z_{\underline{k}'} \delta(\omega - E_{\underline{k}'}) \quad (397)$$

by

$$\frac{Z_{\underline{k}'}^2 \Gamma_{\underline{k}'}}{(\omega - E_{\underline{k}'})^2 + Z_{\underline{k}'}^2 \Gamma_{\underline{k}'}^2 / 4} \quad (398)$$

If we perform the limit $\Gamma \rightarrow 0$ last then we shall see that the ambiguity disappears. The spectral weight function is still sharply peaked so we can put

$$f^-(\omega) = f^-(E_{\underline{k}}) + (\omega - E_{\underline{k}}) \left(\frac{\partial f^-(\omega)}{\partial \omega} \right)_{\omega = E_{\underline{k}}}$$

This gives as $T \rightarrow 0$

$$\begin{aligned} \lim_{q \rightarrow 0} \Lambda_G^I = & - \sum_{\underline{k}'} v(\underline{k} - \underline{k}') \underline{k}' \delta(E_{\underline{k}'} - \mu) \\ & \times \int_{-\infty}^{\infty} \frac{d\omega_1 d\omega_2}{(2\pi)^2} \frac{\omega_1 - \omega_2}{v + \omega_1 - \omega_2} A_{\underline{k}'}(\omega_1) A_{\underline{k}'}(\omega_2) \end{aligned} \quad (399)$$

By substituting the impurity spectral weight function into (399) and performing the integral by contour integration we get

$$\lim_{\Gamma \rightarrow 0} \Lambda_G^I = \sum_{\underline{k}'} v(\underline{k} - \underline{k}') \delta(E_{\underline{k}'} - \mu) \underline{k}' \frac{i Z_{\underline{k}'}^2 \Gamma}{v + i Z_{\underline{k}'} \Gamma} \quad (400)$$

and we see that

$$\lim_{\Gamma \rightarrow 0} \lim_{v \rightarrow 0} \lim_{q \rightarrow 0} \Lambda_G^I = Z_{k_F}^2 \sum_{\underline{k}'} v(\underline{k} - \underline{k}') \delta(\mu - E_{\underline{k}'}') \underline{k}' \quad (401)$$

We have therefore shown that for the particular graph we are considering

$$\lim_{\Gamma \rightarrow 0} \lim_{v \rightarrow 0} \lim_{q \rightarrow 0} \Lambda_G^I = \lim_{q \rightarrow 0} \lim_{v \rightarrow 0} \lim_{\Gamma \rightarrow 0} \Lambda_G^I \quad (402)$$

The ambiguity in the order of limits occurs only in the type of graph we have considered, and Eq. (402) is therefore valid in general. This is an important result since using the reversed order of limits we can forget about the impurities altogether in the vertex function. We shall see later that the ambiguity we have discussed has a simple physical interpretation and is related to the difference between the group and phase velocity of the quasi-particle excitations. In the presence of the impurities we must deal with excitations that are localized between the impurity sites. The whole delicacy of the matter lies in the fact that we are dealing with a current that is carried with the group velocity of the excitations.

To express the vertex function in terms of the parameters of the Landau theory we need any identity similar to the Ward identity in quantum electrodynamics. Identities of this type follow from the conservation laws. The

particular one we shall deduce here follows from charge conservation which implies that

$$\frac{\partial n(\underline{x})}{\partial t} + \nabla \cdot \underline{j}(\underline{x}) = 0 \quad (403)$$

where

$$n(\underline{x}) = e \psi^\dagger(\underline{x}) \psi(\underline{x}) \quad (404)$$

In momentum space

$$n(\underline{q}) = e \sum_{\underline{k}} a_{\underline{k}}^+ a_{\underline{k}+\underline{q}} \quad (405)$$

and Eq. (403) becomes

$$\frac{\partial}{\partial t} n(\underline{q}, t) + i \underline{q} \cdot \underline{j}(\underline{q}, t) = 0 \quad (406)$$

Therefore

$$\begin{aligned} & \frac{\partial}{\partial t'} T < a_{\underline{k}-\underline{q}}(t) n_{\underline{q}}(t) a_{\underline{k}}^+(0) > + i T < a_{\underline{k}-\underline{q}}(t) \underline{q} \cdot \underline{j}(\underline{q}, t) a_{\underline{k}}^+(0) > \\ & = \delta(t'-t) T < [n_{\underline{q}}(t), a_{\underline{k}-\underline{q}}(t)] a_{\underline{k}}^+(0) > \\ & + \delta(t') T < a_{\underline{k}-\underline{q}}(t) [n_{\underline{q}}(0), a_{\underline{k}}^+(0)] > \end{aligned} \quad (407)$$

where the δ functions come from the derivatives of the step functions in the time-ordered product. We had previously shown that

$$\int_0^T dt' \int_0^T dt e^{i(\zeta_l - \nu_m)(t-t')} e^{i\zeta_l t'} T < a_{\underline{k}-\underline{q}}(t) j(\underline{q}, t) a_{\underline{k}}^+(0) >$$

$$= -G_{\underline{k}-\underline{q}}(\zeta_l - \nu_m) \Delta(\underline{k}, \zeta_l; \underline{q}, \nu_m) G_{\underline{k}}(\zeta_l) \quad (408)$$

The limit $\nu \rightarrow 0$ can be performed immediately. Equation (408) suggests that we multiply both sides of (407) by

$$\int_0^T dt' \int_0^T dt e^{i\zeta_l t}$$

In the first term in (407) the integration can be performed trivially, and from the cyclic invariance of the trace it is seen that the contributions from the upper and lower limits of the integration cancel. The commutators on the right-hand side of Eq. (407) are easily computed and we are left with

$$-i G_{\underline{k}-\underline{q}}(\zeta_l) \underline{q} \cdot \underline{\Delta}(\underline{k}, \zeta_l, \underline{q}, 0) G_{\underline{k}}(\zeta_l)$$

$$= -ie G_{\underline{k}}(\zeta_l) + ie G_{\underline{k}-\underline{q}}(\zeta_l) \quad (409)$$

or

$$\underline{q} \cdot \underline{\Delta}(\underline{k}, \zeta_l, \underline{q}, 0) = e G_{\underline{k}-\underline{q}}^{-1}(\zeta_l) - e G_{\underline{k}}^{-1}(\zeta_l) \quad (410)$$

This gives

$$\lim_{q \rightarrow 0} \lim_{\nu \rightarrow 0} \lim_{\Gamma \rightarrow 0} \underline{\Lambda}(\mathbf{k}, \zeta_l) = -e \frac{\partial G_{\mathbf{k}}^{-1}}{\partial \underline{\mathbf{k}}}(\zeta_l) \quad (411)$$

Now

$$G_{\mathbf{k}}^{-1}(\zeta_l) = \zeta_l - \epsilon_{\underline{\mathbf{k}}} - \Sigma_{\mathbf{k}}(\zeta_l) \quad (122)$$

so that

$$\underline{\Lambda}(\mathbf{k}, \zeta_l) = e \frac{\partial \epsilon_{\underline{\mathbf{k}}}}{\partial \underline{\mathbf{k}}} = \frac{e}{Z_{\mathbf{k}_F}} \left. \frac{\partial E_{\underline{\mathbf{k}}}}{\partial \underline{\mathbf{k}}} \right|_{\mathbf{k}_F} \quad (412)$$

Note that $\partial E_{\underline{\mathbf{k}}} / \partial \underline{\mathbf{k}}$ is just the group velocity of the quasi particles. In our previous calculation which did not include the vertex corrections we found

$$\begin{aligned} \frac{e^2}{12 \pi m^2 (\text{Vol})} &< \sum_{\underline{\mathbf{k}} \underline{\mathbf{k}}'} \underline{\mathbf{k}} \cdot \underline{\mathbf{k}}' A_{\underline{\mathbf{k}} \underline{\mathbf{k}}'}(\mu) A_{\underline{\mathbf{k}}' \underline{\mathbf{k}}}(\mu) >_{\text{imp}} \\ &= \frac{ne^2 \tau_{\text{tr}}}{m^*} \left(\frac{m^* Z_{\mathbf{k}_F}}{m} \right)^2 \end{aligned} \quad (357)$$

Substituting this result into the exact expression for the conductivity we get

$$\sigma = ne^2 \tau_{\text{tr}} / m^* \quad (359)$$

which is just the result obtained from the Landau theory using the Boltzmann equation. Although no reference to the Boltzmann equation has been made above, the result (359) suggests that we may have also derived this

equation. In fact it can be shown that the integral equation for the quantity $[\text{Re } K^{++}(k) - K^{+-}(k)]$ of Eqs. (322) - (324) may be interpreted as the Boltzmann equation. We shall not, however, go into this here. It is worth noting, though, that for (359) to be valid we have had only to require that the temperature be low in the sense

$$kT \ll \mu \quad (413)$$

Naively one might have expected that for a kinetic equation to be valid, the lifetimes of the states would have to be long compared to the thermal time \hbar/kT , i. e.,

$$\hbar \Gamma \ll kT \quad (414)$$

We have found, however, that no such condition limits the validity of (359).

Several other remarks are appropriate at this point. In the first place, to bring out the physical significance of the vertex correction, we should mention that by using momentum conservation in the same way as charge conservation was used above, we can derive the identity

$$\lim_{v \rightarrow 0} \lim_{q \rightarrow 0} \lim_{\Gamma \rightarrow 0} \Delta(k_F, \mu) = \frac{ek_F}{m} \cdot \frac{1}{Z_{k_F}} \quad (415)$$

Note that (415) involves the phase velocity at the Fermi surface. If (415) was used in the expression (382) for the conductivity, one would obtain an incorrect result. The essence of the vertex correction is that in the presence of impurities the quasi particles move with their group velocity.

Another point to be noted in (359) is that there is no effective charge correction. An apparent effective charge correction only occurs in the theory when the quasi-particle wave packet is not normalized to one quasi particle per unit volume.

Finally it should be noted that although the effective mass m^* occurs in the result, a d-c conductivity measurement is not really of any use in determining m^* . The reason is that the fundamental quantity is not the scattering time but the cross section which is given by

$$\begin{aligned}\frac{1}{\tau_{tr}} &= \frac{n_i k_F}{m^*} \int \left(\frac{d\sigma_{sc}}{d\Omega} \right) (1 - \cos \theta) d\Omega \\ &= \frac{n_i k_F}{m^*} \sigma'_{sc}\end{aligned}\quad (416)$$

where σ'_{sc} is the transport cross section. We have therefore

$$\sigma(0) = \frac{ne^2}{n_i k_F \sigma'_{sc}} \quad (417)$$

and the effective mass no longer appears. A final comment. If one makes a phase shift analysis of the scattering cross section according to†

$$\frac{d\sigma_{sc}}{d\Omega} = \left| \frac{1}{2ik_F} \sum_{\ell} 2(2\ell+1) P_{\ell}(\cos \theta) \left(e^{2i\delta_{\ell}(k_F)} - 1 \right) \right|^2 \quad (418)$$

one has, using the Friedel sum rule

$$Z = \frac{2}{\pi} \sum_{\ell} (2\ell+1) \delta_{\ell}(k_F) \quad (25)$$

a set of equations that gives us some information about the impurity screening. Equations (417), (418), and (25) have been used for this purpose by Kohn and Vosko.⁽¹⁷⁾ They consider an alloy of a small amount of copper in zinc and assume only s and p scattering. They thus determine δ_0 and δ_1 in terms of the known Z and σ . We see that such an analysis in no way assumes a weak electron-electron interaction.

† We have here reintroduced the spin degree of freedom that was suppressed in the foregoing.

Thermal Conductivity and Concluding Remarks

Another transport property that can be calculated along the same lines as the residual resistivity is the thermal resistance due to impurities. This has also been done by Langer.⁽¹⁸⁾ We shall not discuss this in anything like the detail given for the electrical resistance, but merely sketch how the argument proceeds. In the single-particle approach in Sec. 1 the starting point for a calculation of thermal conductivity was the set of equations (42). One can arrive at the basic structure of these equations from quite general arguments⁽¹⁹⁾ that are independent of the Boltzmann equation. However, a microscopic theory is needed in order to calculate the form of the coefficients L_0 , L_1 , L_2 . We already have an expression for L_0 , since this is just the d-c electrical conductivity, namely,

$$L_0 = \frac{2}{3(\text{Vol})} \text{Im} \int_{-\infty}^{\infty} \langle \underline{j}(0) \cdot \underline{j}(t) \rangle t dt \quad (245)$$

We recall that Eq. (245) was obtained by calculating the electrical current that flows in response to an electric field. In the same way one can calculate the heat current that flows in response to an electric field obtaining

$$L_1 = \frac{2}{3(\text{Vol})} \text{Im} \int_{-\infty}^0 \langle \underline{Q}(0) \cdot \underline{j}(t) \rangle t dt \quad (419)$$

Here \underline{Q} is the quantum-mechanical heat current operator which we shall not specify in greater detail except to say that it is obtained by writing the time derivative of the Hamiltonian density $h(x)$ of our system in the form

$$\frac{\partial}{\partial t} h(x) + \underline{\nabla} \cdot \underline{Q} = 0 \quad (420)$$

From the nature of its derivation it is clear that the expression (419) belongs in the lower left-hand corner of the block of equations (42). That it also belongs in the upper right-hand corner is a consequence of Onsager's reciprocity relations which we here shall simply accept as true. The apparent asymmetry introduced into the set of equations (42) may be removed by using time-reversal invariance. Using this invariance one easily finds that L_1 may also be written in the form

$$L_1 = \frac{2}{3(\text{Vol})} \text{Im} \int_{-\infty}^0 \langle \underline{j}(0) \cdot \underline{Q}(t) \rangle t dt \quad (421)$$

From the form of (245), (419), and (421) it becomes highly suggestive that L_2 has the form

$$L_2 = \frac{2}{3(\text{Vol})} \text{Im} \int_{-\infty}^0 \langle \underline{Q}(0) \cdot \underline{Q}(t) \rangle t dt \quad (422)$$

Although this is not a derivation of the form for L_2 , we are not aware of any that is fundamentally more profound. The result is almost certainly correct, and we shall here accept it without further question. Given L_0 , L_1 , and L_2 we want now to calculate the heat conductivity which we recall is given by

$$\mathcal{H} = \frac{1}{T} \left(L_2 - \frac{L_1^2}{L_0} \right) \quad (45)$$

The thermal conductivity is appreciably more difficult to calculate than the electrical conductivity because it does not suffice to calculate L_0 , L_1 , and L_2 to lowest order in the temperature. To get a nonzero result one has to keep the first temperature corrections (proportional to T^2). These corrections come from two sources in the reduced graph expansion: there is the T^2 contribution of the lowest-order graph, Fig. 22(a), and the contribution of the next reduced graph, Fig. 22(b), to lowest order in the temperature. The T^2 contribution of the lowest graph itself comes from two sources, namely from the explicit Fermi factors in (377) and from the temperature dependence of the electron-electron width Γ . Langer shows that this second temperature dependence is exactly canceled by the part of the T^2 contribution of the reduced graph of Fig. 22(b) that diverges as n_i the density of impurities goes to zero. There remains a contribution from this graph which is finite as $n_i \rightarrow 0$, but it is consistent with the aim of calculating the transport coefficients to order $(n_i)^{-1}$ to neglect this. The basic reason for the cancelation is momentum conservation. Once this cancelation is taken into account, it is shown by Langer that the transport coefficients L_0 , L_1 , L_2 to order

T^2 have the form

$$L_i = - \frac{1}{e} \int dE \frac{\partial f^-}{\partial E} (E, T) E^i \sigma(E) \quad i = 0, 1, 2 \quad (44)$$

From (44) the Wiedemann-Franz law follows as in Sec. 1. It is important to note, however, that for the many-body system this law only holds to order n_i^{-1} .

We have now reached a natural place to end these lectures. Before concluding, however, I should like to make some final remarks. First, it should be mentioned that we have only considered a particularly simple class of transport properties—namely those in which the scattering mechanism is external to the dynamical medium. The Landau theory⁽¹⁾ also envisaged scattering between quasi particles as giving rise, for example, to a finite heat conductivity in a pure system at nonzero temperatures. I am not aware of a detailed derivation of such effects from a microscopic point of view but they seem to be within reach of the formal methods we have discussed.

Finally, I should like to make a comment concerning the general philosophy of the approach that has been taken. We have seen that the use of perturbation theory leads to a theory of the normal Fermi system. The characteristic features of this system are the existence of a many-particle analogue to the Fermi surface, and of low-lying excitations that may be put into one-to-one correspondence with those of the noninteracting system. Now, there is good experimental evidence that nonsuperconducting metals, for example, possess these characteristic properties. Perhaps the most striking evidence of a sharp Fermi surface is the recent observation by Brockhouse⁽²⁰⁾ of the Kohn effect in lead. The effect is an anomaly in the lattice vibration spectrum for wave numbers $q = 2k_F$. The anomaly implies that the ability of the electronic system to screen the ionic charges changes suddenly at $q = 2k_F$. Precisely this behavior would be expected for an electronic system with a sharp Fermi surface because the phase space for energy conserving transitions with momentum transfer q decreases abruptly at $2k_F$. The effect was predicted by Kohn⁽²¹⁾ on this basis.

This and other such strong experimental evidence indicates that some elements of the theory we have been using are strictly correct. It should be emphasized, however, that it would be foolish to believe that the interacting system is normal merely because every order of perturbation theory so predicts. There certainly are fermion systems that do not behave like normal systems, for example, superconductors. Since there is a microscopic theory of superconductivity, namely the BSC theory, one can test to see what goes wrong with perturbation theory in this case. One finds⁽²²⁾ that for fermion systems with attractive interactions, the Green's function

describing the propagation of a pair of particles acquires, in perturbation theory, a pole on the imaginary axis of the physical part of the ζ plane. Such a pole is in contradiction to the spectral representation and leads to an amplitude for the propagator that increases exponentially with time. This unphysical behavior reflects the instability of the normal system with respect to the formation of Cooper pairs. (23) It may be a general feature that if such instabilities are not present, the perturbation theory leads to sensible results, and it may be that the repulsive Coulomb interaction does not give rise to such instabilities. Even if it does, however, it is likely that those analytic properties of a general order of perturbation theory that we have used in our discussion are preserved in a certain class of nonperturbative solutions. These questions are all difficult to answer, and we leave them for future lecturers at this hospitable summer school.

APPENDIX A: LOW-TEMPERATURE FORM OF $\Gamma_k(\omega)$

In this appendix we discuss the low-temperature form of the electron self-energy operator and show that

$$\lim_{T \rightarrow 0} \lim_{\omega \rightarrow \mu} \Gamma_k(\omega) = 2C_k \left[(\omega - \mu)^2 + \pi^2 k_B^2 T^2 \right] + \dots \quad (\text{A-1})$$

The easiest way to prove (A-1) is to realize that $\Gamma_k(\omega)$ is the discontinuity across the cut of $\Sigma(\zeta)$, and can, therefore, be evaluated in terms of reduced graphs. Figure A.1 shows the two lowest-order graphs.



Fig. A.1

According to our discussion of reduced graphs in Sec. 4, the contribution of Fig. A.1 (a) is

$$\Gamma^{(1)}(\omega) = \sum_{\text{mom}} \int \frac{d\omega_1 d\omega_2 d\omega_3}{(2\pi)^3} A(\omega_1) A(\omega_2) A(\omega_3) \quad (\text{A-2})$$

$$\times |\Lambda(\omega - i\eta)|^2 \frac{f^+(\omega_1) f^+(\omega_2) f^-(\omega_3)}{f^+(\omega)} 2\pi \delta(\omega + \omega_3 - \omega_1 - \omega_2)$$

Note that this is an implicit nonlinear equation for Γ since the A 's depend on Γ according to Eq. (127). Note that at low temperatures and for $\omega \geq \mu$ the integrand is strongly peaked near $\omega_1, \omega_2, \omega_3 \sim \mu$ because of the Fermi factors. We assume, and later verify that the rest of the integrand when summed over the momenta is perfectly smooth near this point. Then we find in leading order

$$\Gamma^{(1)}(\omega) \propto \frac{1}{f^+(\omega)} \int \frac{d\omega_1 d\omega_2 d\omega_3}{(2\pi)^3} f^+(\omega_1) f^+(\omega_2) f^-(\omega_3) \quad (\text{A-3})$$

$$\times 2\pi \delta(\omega + \omega_3 - \omega_1 - \omega_2)$$

The integral (A-3) can be done explicitly⁽²⁴⁾ leading to the result (A-1). At $T = 0$ the integral is trivial since for $\omega \geq \mu$ we have

$$\Gamma^{(1)}(\omega) \propto \int_{\mu}^{\infty} \frac{d\omega_1}{2\pi} \int_{\mu}^{\infty} \frac{d\omega_2}{2\pi} \int_{-\infty}^{\omega} \frac{d\omega_3}{2\pi} 2\pi \delta(\omega + \omega_3 - \omega_1 - \omega_2) \quad (\text{A-4})$$

One of the integrals is removed by the δ function. What remains is the area of a small triangle in the space of the two remaining variables. This is proportional to $(\omega - \mu)^2$. For $\omega \leq \mu$ we notice that, for $\omega = \omega_1 + \omega_2 - \omega_3$,

$$\frac{f^+(\omega_1) f^+(\omega_2) f^-(\omega_3)}{f^+(\omega)} = \frac{f^-(\omega_1) f^-(\omega_2) f^+(\omega_3)}{f^-(\omega)} \quad (\text{A-5})$$

Then the above argument can be repeated leading to the same result.

Now it is easy to verify the assumption of smoothness for the rest of the integrand in (A-2). We already saw in Sec. 2 that the form (A-1) for Γ leads to at most a δ function in A at the Fermi surface. The integration over momentum, however, removes this singular behavior. The assumption is therefore justified.

In the same way it can be shown that the graph (b) of Fig. A-2 gives a contribution to Γ at $T = 0$ that is proportional to $(w - \mu)^4$. The leading dependence is thus given by the graph we have considered.

APPENDIX B: LUTTINGER'S THEOREM

In Sec. 3 use was made of the identity

$$X = \lim_{T \rightarrow 0} \frac{1}{\beta} \sum_{\zeta_\ell, p_\ell} G_p(\zeta_\ell) \frac{\partial}{\partial \zeta_\ell} \sum_p (\zeta) = 0 \quad (\text{B-1})$$

We give here a proof of this identity. Because the spacing of the points ζ_ℓ is $2\pi i/\beta$, we can replace the left side of (B-1) by

$$\frac{1}{2\pi i} P \int_{\mu-i\infty}^{\mu+i\infty} \sum_p G_p(\zeta) \frac{\partial}{\partial \zeta} \sum_p (\zeta) d\zeta \quad (\text{B-2})$$

where P means the principal value at $\mu = 0$. In the most general case the integrand of (B-2) is not a smooth function in the neighborhood of $\zeta = \mu$ because, as we have seen, the imaginary part of $\Sigma(\zeta)$ has a discontinuity Γ across the real axis leading to a δ function-like contribution to $\partial \Sigma / \partial \zeta$. However, for the system we are considering $\Gamma(\mu) = 0$ at $T = 0$ as was shown in Appendix A. Because of this we may remove the principal value sign without changing the value of the integral. Then, doing an integration by parts one finds

$$X = - \frac{1}{2\pi i} \int_{\mu-i\infty}^{\mu+i\infty} \sum_p \frac{\partial G_p(\zeta)}{\partial \zeta} \sum_p (\zeta) d\zeta \quad (\text{B-3})$$

As written in the last line X has an immediate graphical expansion. Consider the graphs for the quantity Ω' of Eq. (197). If the Green's function corresponding to every line of a graph for Ω' is differentiated in turn with respect to the corresponding quasi energy and then each summation $(1/\beta) \times \sum_{\zeta_l}$ is replaced by an integration

$$\frac{1}{2\pi i} \int_{\mu-i\infty}^{\mu+i\infty} d\zeta$$

a term in the expansion of X results. Summing over all graphs for Ω' then gives X . However, the quasi-energies for each line are connected by conservation conditions. Doing partial integrations in the terms for X one finds factors of the form

$$\left(\frac{\partial}{\partial \zeta_1} + \frac{\partial}{\partial \zeta_2} + \frac{\partial}{\partial \zeta_3} + \frac{\partial}{\partial \zeta_4} \right) \delta(\zeta_1 + \zeta_2 - \zeta_3 - \zeta_4) = 0 \quad (\text{B-4})$$

Therefore X is zero.

APPENDIX C: OPTICAL SUM RULE

In this appendix we show that the coefficient of $1/\omega$ in the expression (244) for the conductivity is in fact zero for the system considered. We need to prove that†

$$\frac{1}{3(V\omega l)} \int_{-\infty}^0 < \underline{J}(0) \cdot \underline{J}(t) - \underline{J}(t) \cdot \underline{J}(0) > dt = \frac{ne^2}{m} \quad (\text{C-1})$$

† Note that (C-1) is certainly not true in the absence of impurities. In that case, the total momentum P is a constant of the motion, and since $J \propto P$ and the thermal ensemble has no net momentum, the left-hand side of (C-1) is 0. $T = 0$ where T is a large time.

In terms of the quantities $F(t)$ and φ introduced in Eq. (246) and (252) the left-hand side of (C-1) is

$$\frac{1}{3(Vol)} \int_{-\infty}^0 (F^{<}(t) - F^{>}(t)) dt = \frac{1}{3(Vol)} \int_{-\infty}^{\infty} \frac{\varphi(\omega)}{\omega} \times \frac{d\omega}{2\pi} = \frac{i}{2(Vol)} F(v_m = 0) \quad (C-2)$$

The last identity follows from (252). We may evaluate (C-2) as follows

$$F(t) = \frac{-e}{m} T \int \left(\frac{\nabla_1 - \nabla_1'}{2i} \right) < \psi(xt) \underline{J}(0) \psi^+(xt^+) > d^3x \\ = \frac{-e}{m} \sum_{\underline{k}} \underline{k} \cdot T < a_{\underline{k}}(t) \underline{J}(0) a_{\underline{k}}^+(t^+) > \quad (C-3)$$

Evaluating the Fourier series coefficient in the restricted time region we have

$$F(v_m) = \frac{e}{m\tau} \sum_{\underline{k}, \zeta_l} \underline{k} \cdot G_{\underline{k}}(\zeta_l + v_m) \underline{\Lambda}(\underline{k}, \zeta_l; q=0, v_m) \\ \times G_{\underline{k}}(\zeta_l) e^{\zeta_l 0^+} \quad (C-4)$$

where $\underline{\Lambda}$ is the vertex function introduced in Eq. (377). We have already seen that for the system with impurities

$$\lim_{v \rightarrow 0} \underline{\Lambda}(\underline{k}, \zeta_l, q=0, v_m) = -e \frac{\partial}{\partial \underline{k}} G^{-1}(\underline{k}, \zeta_l) \quad (411)$$

Thus

$$\begin{aligned}
 \frac{i}{3(\text{Vol})} F(\nu_m = 0) &= -\frac{ie^2}{3m} \frac{1}{\tau(\text{Vol})} \sum_{\mathbf{k}, \zeta_l} \underline{\mathbf{k}} \cdot \frac{\partial}{\partial \underline{\mathbf{k}}} \\
 &\times G(\mathbf{k}, \zeta_l) e^{\zeta_l 0^+} = +\frac{ie^2}{m} \frac{1}{\tau(\text{Vol})} \\
 &\times \sum_{\zeta_l} G(\mathbf{k}, \zeta_l) e^{\zeta_l 0^+} = \frac{ne^2}{m}
 \end{aligned} \tag{C-5}$$

which proves the desired identity. The significance of this equality may be seen by noting that from (244), (246), and (252)

$$\text{Re } \sigma(\omega) = \frac{1}{6(\text{Vol})} \frac{\varphi(\omega)}{\omega} \tag{C-6}$$

Thus (C-5) expresses the sum rule

$$\text{Re} \int_{-\infty}^{\infty} \sigma(\omega) d\omega = \frac{ne^2 \pi}{m} \tag{C-7}$$

One can argue quite generally that this sum rule must be satisfied for a dissipative system. The argument is as follows. It follows from causality that $\sigma(\omega)$ considered as a function of complex ω is analytic in the upper half ω plane. Furthermore the asymptotic behavior of $\sigma(\omega)$ is determined by the density, charge, and real mass of the particles in the system according to

$$\lim_{\omega \rightarrow \infty} \sigma(\omega) = \frac{-ne^2}{mi\omega} \tag{C-8}$$

Therefore we can write a dispersion relation for $\omega \sigma(\omega)$ using Cauchy's theorem, namely,

$$2\pi i \omega \sigma(\omega) = \int_{-\infty}^{\infty} \frac{\omega' \sigma(\omega')}{\omega' - \omega} d\omega' - \frac{ne^2 \pi}{m} \quad (\text{C-9})$$

where ω is in the upper half-plane. Now for a dissipative system

$$\lim_{\omega \rightarrow 0} \omega \sigma(\omega) = 0 \quad (\text{C-10})$$

By using (C-10) in (C-9) we obtain the sum rule (C-7).

APPENDIX D: EFFECTIVE MASS SUM RULE

We give here a proof of the effective mass sum rule (23). In the course of doing this we shall identify the interaction function of the Landau theory which we called $\tilde{\Gamma}$ in Sec. 1. We follow the method of Luttinger and Nozières.⁽²⁵⁾

Consider the vertex function defined according to

$$(-i)^2 T \langle \psi(1) \bar{\psi}(2) \psi^\dagger(1') \rangle = G(1 \bar{1}) \Lambda(\bar{1} \bar{1}'; 2) G(\bar{1}' 2) \quad (\text{D-1})$$

If we further define the kernel Γ in terms of the two-particle Green's function as

$$\begin{aligned} G(1 2; 1' 2') &= G(1 1') G(2 2') - G(1 2') G(2 1') \\ &+ i G(1 \bar{1}) G(2 \bar{1}) \Gamma(\bar{1} \bar{1}'; \bar{2} \bar{2}') \\ &\times G(\bar{2} 2') G(\bar{2}' 2') \end{aligned} \quad (\text{D-2})$$

then

$$\begin{aligned} \underline{\Lambda}(1 \ 1'; 2) &= \frac{\nabla_2 - \nabla_{2'}}{2 \ i} \left[\delta(1 \ 2') \delta(1' 2') \right]_{2' = 2} \\ &\quad - i \frac{\nabla_2 - \nabla_{2'}}{2 \ i} G(2 \ 2) \Gamma(1 \ 2; 1' 2') \\ &\quad \times G(2' 2') \Big|_{2' = 2^+} \end{aligned} \quad (D-3)$$

Fourier analyzing this equation in the restricted time region gives

$$\begin{aligned} \underline{\Lambda} \left(\underline{k}, \zeta, \underline{q}, \nu; \underline{m} \right) &= \left(\underline{k} + \frac{\underline{q}}{2} \right) \cdot \frac{-i}{\tau} \sum_{\underline{k}', \zeta'} \left(\frac{\underline{k}'}{2} + \frac{\underline{q}'}{2} \right) \\ &\quad \times \Gamma(\underline{k}, \zeta; \underline{k}', \zeta'; \underline{q}, \nu) G(\underline{k}', \zeta') G(\underline{k}' + \underline{q}', \zeta' + \nu) \end{aligned} \quad (D-4)$$

Graphically Eq. (D-4) is represented by Fig. D.1.

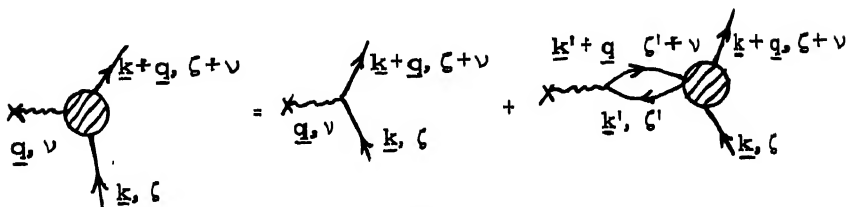


Fig. D.1

Now one can write an integral equation for Γ which is represented by Fig. D.2. Using a four-dimensional notation $\underline{k} = (\underline{k}, \zeta)$ this equation reads

$$\begin{aligned} \Gamma_{\underline{k} \underline{k}'}(\underline{q}) &= I_{\underline{k} \underline{k}'}(\underline{q}) + \frac{i}{\tau} \sum_{\underline{k}''} I_{\underline{k} \underline{k}''}(\underline{q}) \\ &\quad \times G_{\underline{k}''} G_{\underline{k}'' + \underline{q}} \Gamma_{\underline{k}' \underline{k}'}(\underline{q}) \end{aligned} \quad (D-5)$$

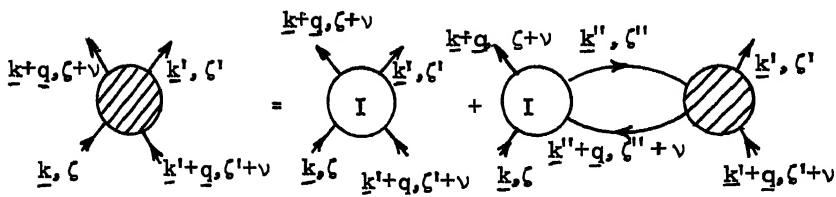


Fig. D.2

Above $I(q)$ is the irreducible part of the kernel, i.e., the part that cannot be broken into two parts by cutting two lines. We have already seen in the discussion of vertex functions in Sec. 4 that the contribution of a single electron-hole pair excitation has a discontinuous behavior for small total momentum and energy. Writing

$$R_{\mathbf{k}\mathbf{k}'}(q) = \frac{i}{\tau} \delta_{\mathbf{k}\mathbf{k}'} G_{\mathbf{k}} G_{\mathbf{k}+q} \quad (\text{D-6})$$

we may extract the discontinuous part of R at $T = 0$ as indicated in Eqs. (391) to (394) and write

$$R_{\mathbf{k}\mathbf{k}'}(q) = R_{\mathbf{k}\mathbf{k}'} + \Delta_{\mathbf{k}\mathbf{k}'}(q) \quad (\text{D-7})$$

where

$$\Delta_{\mathbf{k}\mathbf{k}'}(q) = Z_{\mathbf{k}}^2 \frac{\mathbf{q} \cdot \mathbf{V}_{\mathbf{k}}}{\mathbf{q} \cdot \mathbf{V}_{\mathbf{k}} - v} \delta(E_{\mathbf{k}} - \mu) \delta(\zeta - \mu) \quad (\text{D-8})$$

Here $\mathbf{V}_{\mathbf{k}} = \partial E_{\mathbf{k}} / \partial \mathbf{k}$. We expect that the remainder will be a smooth function for small q and v . Substituting (D-6) into the integral equation (D-5) we get in a matrix notation

$$\Gamma(q) = I(q) + I(q) R(q) \Gamma(q) \quad (\text{D-9})$$

We consider now the limit $q = 0$, $\nu \rightarrow 0$. Since Δ is zero in this limit we may write (D-9) in the form

$$\Gamma(q) = \Gamma^w + \Gamma^w \Delta(q) \Gamma(q) \quad (D-10)$$

where the superscript indicates the above-mentioned limit. The integral equation (D-4) for Λ may be written in the form

$$\Lambda(k, q) = \underline{k} - \sum_{k''} \Gamma_{kk''}(q) R_{k''k'}(q) \underline{k'} \quad (D-11)$$

If we substitute Γ in (D-11) from (D-10) and note that

$$R = R^w + \Delta(q) \quad (D-12)$$

[since $\Delta^w(q) = 0$] we find

$$\Lambda(k, q) = \Lambda^w(k) - \Gamma^w \Delta(q) (\underline{k} - \underline{k'}) - \Gamma^w \Delta(q) \underline{k'} \quad (D-13)$$

The last two terms cancel. Now if we take the q limit ($\nu = 0$, $q \rightarrow 0$) of (D-13) we obtain

$$\begin{aligned} \underline{\Lambda}^q(\underline{k}, \zeta) &= \underline{\Lambda}^w(\underline{k}, \zeta) + \sum_{k''} \Gamma^w(k, \zeta, k'' \zeta'') Z_{k''}^2 \delta(E_{k''} - \mu) \\ &\times \underline{\Lambda}^q(k'', \mu) \end{aligned} \quad (D-14)$$

we already saw that for $|k| = k_F$ [cf. Eq. (412)]

$$\underline{\Lambda}^q(\underline{k}, \mu) = \frac{m}{Z_{k_F}} \frac{\partial E}{\partial \underline{k}} \quad (D-15)$$

and it was stated in the text (the proof⁽²⁵⁾ uses the Ward identity corresponding to momentum conservation) that

$$\underline{\Lambda}^W(\underline{k}, \mu) = \frac{1}{Z_{\underline{k}_F}} \underline{k} \quad (\text{D-16})$$

Thus if we identify

$$\bar{\Phi}_{\underline{k}\underline{k}'} = (\text{Vol}) Z_{\underline{k}_F}^2 \Gamma^W(\underline{k}, \mu; \underline{k}', \mu) \quad (\text{D-17})$$

for $|\underline{k}| = |\underline{k}'| = k_F$, Eq. (D-14) reduces to

$$\frac{\underline{k}}{m^*} = \frac{\underline{k}}{m} + \sum_{\underline{k}''} \bar{\Phi}_{\underline{k}\underline{k}''} \frac{\underline{k}''}{m^*} \delta(E_{\underline{k}''} - \mu) \quad (\text{D-18})$$

Doing the radial integral we get

$$\frac{1}{m^*} = \frac{1}{m} + \frac{k_F}{(2\pi)^3} \int d\Omega \cos \theta \bar{\Phi}_{\underline{k}_F}(\theta)$$

This is precisely the effective mass sum rule (23).

Acknowledgments

The preparation of these notes was supported in part by the Office of Naval Research. I am also grateful to Mr. A. Baratoff for some editorial assistance.

References

1. L. D. Landau, Soviet Phys. JETP (English Transl.), 3, 920 (1957); *ibid.* 5, 101 (1957). The microscopic approach was also initiated by Landau in the following reference: Soviet Phys. JETP (English Transl.), 8, 70 (1959).
2. P. C. Martin and J. Schwinger, Phys. Rev., 115, 1342 (1959). See also L. P. Kadanoff and G. Baym "Quantum Statistical Mechanics," Benjamin, New York, 1962.
3. J. M. Luttinger, Phys. Rev., 119, 1153 (1960); J. M. Luttinger and J. C. Ward, Phys. Rev., 118, 1417 (1960).
4. J. S. Langer, Phys. Rev., 120, 714 (1960); Phys. Rev., 124, 1003 (1961).
5. L. D. Landau, Soviet Phys. JETP (English Transl.), 7, 262 (1958).
6. V. Galitski and A. Migdal, Soviet Phys. JETP (English Transl.), 7, 96 (1958).
7. G. Baym and N. Mermin, J. Math. Phys., 2, 232 (1961).
8. J. M. Luttinger, Phys. Rev., 121, 942 (1961).
9. C. Bloch and C. De Dominicis, Nuclear Phys., 7, 459 (1958).
10. J. S. Langer and V. Ambegaokar, Phys. Rev., 121, 1090 (1961).
11. T. D. Lee and C. N. Yang, Phys. Rev., 117, 22 (1960).
12. C. Bloch and C. De Dominicis, Nuclear Phys., 27, 294 (1962).
13. R. Kubo, J. Phys. Soc. Japan, 12, 570 (1957).
14. J. S. Langer, Phys. Rev., 124, 997 (1961); Phys. Rev., 127, 5 (1962).
15. L. D. Landau, Nuclear Phys., 13, 181 (1959).
16. R. D. Cutkosky, J. Math. Phys., 1, 429 (1960).
17. W. Kohn and S. Vosko, Phys. Rev., 119, 912 (1960).
18. J. S. Langer, Phys. Rev., 128, 110 (1962).
19. H. B. Callen, "Thermodynamics," Wiley, New York, (1960)
20. B. N. Brockhouse et al., Phys. Rev. Letters, 7, 93 (1961); Phys. Rev., to be published.
21. W. Kohn, Phys. Rev. Letters, 2, 393 (1959).
22. L. P. Kadanoff and P. C. Martin, Phys. Rev., 124, 670 (1961).
23. L. Cooper, Phys. Rev., 104, 1189 (1956).